OSIS Users Manual

(OSIS Schema 2.1.1)

By

Patrick Durusau

Patrick@Durusau.net

On Behalf of the American Bible Society

Foreword

The OSIS Users Manual is the result of years of effort by organizations and individuals who have generously donated their support and time to taking the delivery of the Bible into the 21st century. Its subject is the use of the OSIS schema, a means to put Bibles into the lingua franca of modern computers, XML (Extensible Markup Language). Once a Bible is in XML, effective delivery in whole or part in a variety of formats (print, WWW, cellphones, etc.) and use of Bible portions in a variety of contexts becomes possible for all organizations and individuals.

This users manual is a guide to the 2.1.1 version of the OSIS schema. Be aware that the OSIS schema continues to evolve and work on OSIS 2.5 is currently underway. Future changes to the schema will be backwards compatible so any documents prepared following this guide will continue to be valid in the future.

The OSIS development team closely studied previous Bible encoding forms, as well as tools for literary encoding in general. By doing this we hope we have avoided some weaknesses, and gained from some strengths, of each one, and we thank the many people who worked on those prior specifications, as well as those who have provided help and feedback in developing OSIS itself, and testing it by encoding large numbers of Biblical and related texts.

Users familiar with the Text Encoding Initiative will find OSIS markup quite familiar, because the bulk of the elements we define correspond directly to TEI elements, and almost always have the same name (though often simplified). The schema also provides a TEIform attribute for such elements, so they can be recognized by form-aware processors as equivalent to their TEI counterparts.

OSIS is provided as a free resource by the Bible Technologies Group™ (or BTG™), which is a collaborative effort of the American Bible Society, the Society of Biblical Literature, the Summer Institute of Linguistics, the United Bible Societies, other Bible Societies and related groups, and individual volunteers around the world. OSIS was designed to meet the needs of diverse user communities who read, study, research, translate or distribute biblical texts.

For more information on OSIS, join the OSIS Users' Group. To do so, send mail to osis-user@whi.wts.edu, setting the Subject line to "subscribe". Online information about OSIS is also available at http://www.bibletechnologies.org and http://www.bibletechnologieswg.org.

The OSIS (Open Scriptural Information Standard) project began as at breakfast meeting organized by Dennis Drescher (SIL) at the XML 2000 meeting in Washington, D.C. Unlike many such meetings about an encoding standard for Bibles and related literature, this one resulted in a coalition that produced both the OSIS schema and the following documentation.

There have been meetings, teleconferences and email discussions numbering in the thousands of posts as each part of the OSIS schema was developed and tested against the needs of those who are interested in using XML in tasks such as publishing, translating and studying the Bible. Work on additional capabilities for OSIS and more documentation is ongoing and newcomers are welcome to the effort. Comments, corrections, supplied omissions/examples and questions about participating should be addressed to to osis-editors@bibletechnologieswg.org.

Over the long process to this point, the OSIS project has been supported by the American Bible Society (special thanks to Dr. Eugene Habecker, and Rev. Trevon Gross), CrossWire.org, SIL, the Society of Biblical Liteature (special thanks to Dr. Kent Harold Richards), Snowfall Software, and the United Bible Societies. These organizations have contributed staff, resources, organized meetings, paid teleconference bills and supported in other ways the development of OSIS. Suffice it to say that OSIS would not have been possible but for their support.

Organizations played an important role in this effort, but then so did a very large number of individuals. Any listing of individuals will be incomplete but it would be remiss to not name some of the key figures who are responsible for the schema and documentation you see today:

- Eric Albright: Eric spent several days at the first OSIS meeting in Rome enlightening the OSIS core group on the choices and decisions made in the creation of XSEM. Eric's insight into the problems faced by translators is profound and was instrumental in shaping core features of the OSIS schema.
- Jim Albright: The common last name with Eric Albright is no accident as Jim is Eric's father. Jim has commented extensively on the various drafts of OSIS and saved the reader from both substantive as well as typographical errors on any number of occasions.
- Jeff Bone: Jeff was the head of ABSi and played a major role in pledging ABS involvement in OSIS.
- Alan Connor: My definition of a great manager. Rather than look for paperwork or forms, Alan has always looked for solutions. Alan's friendship and assistance in OSIS meetings at SIL have been deeply appreciated.
- Peter Constable: Peter has been deeply involved in enabling people to speak with their own voices as part of the Non-Roman Scripts Initiative at SIL for many years. We have valued Peter's advice on character set issues so that OSIS will continue that tradition.
- Kees DeBlois: Kees is the long time head of technical development at UBS and OSIS was quite fortunate when Kees agreed to serve as co-chair with Steve DeRose. Kees contributed a much needed "translators" view to the encoding practices found in the OSIS schema.
- Dennis Drescher: The organizer of the meeting that gave birth to the the OSIS initiative. Dennis has been a long time supporter of OSIS and a participant in our discussions.
- Steve DeRose: Steve agreed to become the co-chair of the Bible Technologies Group (the 'official' activity that lead to OSIS) along with Kees DeBlois in the Spring of 2001. The author of more international standards that can be easily listed, Steve has been both a source of technical work as well as keeping the more excitable members of the core group under control.
- Jim De Vries: Jim has participated in several OSIS meetings at SIL and along with Darrell Eppler, helped the OSIS team become aware of publication issues. Simply not enough to be elegant, it has to result in printed output as well.
- John Edwards: John's proof reading of one of the drafts of this manual made me wonder if I had read any of it while I was typing it. If the following manual is readable at all, it is due to the efforts of John and others to filter the output of my keyboard.
- Daniel Englbauer: Daniel was one of several international participants in the OSIS project and an able errata contributor.
- Darrell Eppler: Darrell has helped OSIS to be mindful of the need to go from some file on a computer system to something that the average reader would recognize as a Bible. Abstract markup is all well and good, but at some point, some people anyway, want a physical object called a Bible in their hands. To the extent that is possibe, Darrell and Jim Vevries are responsible for that being possible.
- Daniel Glassey: Daniel is one of the many Crosswire volunteers who found time to contribute to the OSIS project.
- Troy Griffitts: Founder of Crosswire.org, one of the largest community projects for developing free Bible software. Troy was one of the early adopters of OSIS and has tirelessly worked on numerous drafts and test conversions of data into proposed OSIS markup.

- Adina Hamik: Adina was an early booster for OSIS within the ABS and helped organize OSIS meetings. Her wit and charm helped keep the various organizations and individuals going in a common direction.
- Bob Hodgson: Bob has been involved in OSIS since the early days of 2001. He wears so many hats at the ABS that it is difficult to describe which one he had on while supporting OSIS. I think it was the general, support Bible engagement one but I can't say for sure.
- Michael Paul Johnson: Michael has influenced the OSIS group on a number of issues and illustrates how participation, even from a distant location can make a real difference in public projects.
- Chris Little: Chris is also involved in Crosswire.org (along with Troy Griffitts) and has been a long term participant in OSIS. Chris is one of the calmer members of the core team, a much needed presence more often than you would think. He has contributed to both testing and substantive development of the schema.
- Kirk Lowery: Kirk is a Hebrew linguistics expert who has been helping to shape OSIS so that linguistic markup can be added to OSIS texts. Almost there Kirk! I promise!
- Nathan Miles: A very good friend and colleague who is charged with not only supporting new formats, such as OSIS, but also supporting old formats as well. A gifted programmer whose suggestions have shaped and refined OSIS.
- Mike Perez: Mike was the critical person in OSIS moving from a breakfast conversation to the schema and documentation that exists today. Mike organized and patiently attended every OSIS meeting and teleconference, while the markup specialists argued about issues too obscure to be repeated here.
- Harry Plantinga: Harry developed ThML, a language for encoding theological texts but found the time to convert materials into OSIS, to chair OSIS meetings, developed an OSIS editing tool and is now converted ThML texts into OSIS.
- Jim Schaad: Jim was an early and frequent errata contributor. Any errors that remain were not due to a lack of effort on his part.
- DM Smith: A reader of the last draft of this manual who saved readers from many confusing and unlikely to be correct statements.
- Todd Tillinghast: Todd writes the stylesheets that enable OSIS documents to suddenly appear like something that an average user might want to read. Among other things, Todd has created numerous test encodings and other material that has assisted in the development of OSIS.
- John Walter: If others served in the trenches of writing encoding schemas and documentation, John served an equally important role of advocate for OSIS both within the ABS and to the larger Bible community. It is due in no small part to his efforts that OSIS is now being adopted by the ABS in many of its activities.

Claiming author's privilege in a foreword, I would like to publicly thank all of the participants in the OSIS project for their time, efforts and faith that has brought OSIS to this point. In a time when faith is often treated with disdain, it has been my privilege to work in this project with men and women of great faith.

Patrick Durusau
OSIS Technical Lead
Covington, Georgia, March 2006

Contents

1. Introduction to OSIS™

Welcome to the OSIS (Open Scriptural Information Standard™) User's Manual. OSIS is an XML schema that can be used to produce Bibles, commentaries, and related texts that can be easily interchanged with other users, formatted as HTML, PDF, Postscript or any other desired format, and searched on any personal computer. It provides a standard way to express such documents, which is important because it saves time, money.

- Authors can prepare one manuscript for different potential publishers.
- Publishers receiving OSIS documents can use standard XML tools to convert to their internal formats, or to go directly to print or electronic products (without the need for further conversion).
- Software vendors can use existing XML software for their products.
- Users can interchange and use Bibles and other texts using software that is freely available to all users.

XML (Extensible Markup Language) may be unfamiliar to some OSIS users and the following is a brief introduction to XML and some of the terms that will be encountered in this manual.

XML is a **descriptive** markup language as opposed to being a **prescriptive** markup language. A **prescriptive** markup language would mark the beginning of a paragraph, for example, with an instruction such as *double line return, indent 5 spaces* in a language understood by its processing engine. The problem is that every processing engine uses a different set of instructions.

A **descriptive** markup language, on the other hand, would mark the beginning of a paragraph with **<p>** and the end of the paragraph with **</p>**. How the paragraph is to be formatted is left up to latter processing. One advantage of making the markup descriptive is that the same file can be used to create HTML, PDF or files for delivery to a cellphone. Another major advantage is that files can be easily exchanged between different processors since the processor specific information is not in the file.

This manual uses the terms **element** and **attribute** quite frequently. An **element** is a portion of text of a particular kind. For example, a **<p>** element represents a different structure in a text than a **<q>** element. Those elements are represented in an XML document by **tags,** that you will see as: **<p>...</p>** in pairs or sometimes alone, **<verse/>** (also known as a **milestone** because it may not have a matching closing tag. The elements permitted in an encoded document (one with XML markup, that is elements and attributes) and which elements can be found inside other elements are specified by a set of rules. That set of rules for OSIS is specified by an XML schema, the OSIS schema.

An **attribute** is additional information that should be recorded but not as part of the text. For example, when marking a quotation from Jesus, the **q** element would be written: **<q who="Jesus">Eli, Eli, lama sabach'thani?</q>**, where the **who="Jesus"** is an attibuted that identifies the quoted material as being said by Jesus. Attributes enable searching for or special presentation of portions of a text, such as rendering all the words of Jesus in red.

The manual is organized as follows:

- **Introduction to OSIS** (this chapter), Getting started, Some authoring tools, Your First OSIS Document. Introductory materials meant for users new to XML and OSIS.
- **XML and OSIS declarations**. XML and OSIS declarations are a "cut-n-paste" proposition for all but the most technically skilled users.

- **The OSIS text header**. To have a useful electronic text, certain information should be recorded about the text that is not really part of the text itself. The OSIS text header is the location of that information in an OSIS document.
- **Work**. Work in this context refers to the text that is being encoded as an OSIS document, which requires "work" in a different sense of the word. Although the material is fairly complex, there are some simple examples given which should be followed until the user develops the background to fully use the more complex aspects of the work element.
- **Date Formats**. In order to effectively search electronic documents for dates, it is necessary for dates to be identified and recorded in some uniform manner. This chapter covers the OSIS rules for dates.
- **Title Pages**. Strictly speaking title pages are composed of information found elsewhere in a document but OSIS provides the ability to encode a title page for those more comfortable with that approach.
- **Div Elements**. The basic structure of an OSIS document starts with dividing the text into large parts, called **div** elements. These elements are used to organize materials into chapters, books, testaments, annexes and similar large portions of text.
- **Paragraphs, Quotes, and Notes**. OSIS rules for paragraphs, quotes, notes and illustrations are covered in this chapter. Many modern translations use paragraphs as opposed to verses to organize the biblical text so users should pay particular attention to this chapter.
- **Interim Summary**. A short listing of the main elements encountered at this point, which illustrates how using only a few elements can produce a useful encoding of a Bible.
- **Elements that cross other elements**. It is often the case that a quotation may start in one verse or paragraph and end in another. OSIS has developed a new technique for handling such cases that will easily handle such cases.
- **Special Text Types**. Texts such as epistles, dramatic texts, speeches and speakers, poetic material, tables, genealogies, and figures (illustrations) are covered in this section.
- **Milestones**. A milestone, like markers along the side of the road, are used to mark particular places in a text. Not every text will have them but for some purposes, such as marking required line breaks, they are essential.
- **Common elements in all texts**. There are a number of things that occur in all texts, such as abbreviations, foreign words, references, names, and others. The markup for these and more are gathered here for easy reference.
- **Bible Editions**. Bibles are often produced in editions that differ in their content. Some will have pronunciation helps, others extensive maps and other features. It is important to distinguish between such editions and the OSIS mechanism for doing so is described in this section.
- **OSIS Identifiers**. The reason why all members of a congregation can read along with readings on Sunday is because they all share one way to refer to the Bible. But, there are many such reference systems and OSIS provides the means to represent the systems expected by any given user community.
- **OSIS References**. It is very common, particularly in study notes and other places, to want to make references to other parts of a Bible, ranges of verses, individual verses, or even parts of verses. OSIS provides techniques for doing all three.
- **Different Versification Systems**. While not encountered by the average user, different versification systems exists for the biblical text. Some Bibles provide references or even text that uses different systems of versification. How that should be encoded in OSIS is covered in this section.
- **Appendices**. The appendices contain reference materials that OSIS users will find helpful. Among other things, there are quick charts of OSIS elements and their attributes, standard abbreviations for books of the Bible and other material. Of particular interest to current Bible translation efforts is the mapping of USFM codes to OSIS elements along with brief notes on their application.

If that sounds like a lot of material, take heart! OSIS was designed to have approximately the same number of elements that were found in HTML 3.2, one of the most successful markup languages ever written. Aside

from the header portion of a text, which may be fairly complex, the average user can produce a professional quality encoding of a Bible text with fifteen (15) or fewer elements. A Bible text that can then be used by other users, publishers, or others, to produce a wide range of materials. Make no mistake, using OSIS does require effort but the payoff in terms of making the Bible universally accessible are immense!

2. Getting started

The first question that is often asked when learning that OSIS uses XML (a markup language) is: 'I'm not a computer person. Can I learn to use OSIS?' If you can type and use even the most basic word processor or computer text-editing program, the answer is clearly 'Yes!' OSIS was designed to offer the beginning user a simple way to do the basic 'markup' required for a standard biblical text. 'Markup' refers to markers placed within the text, that indicate where useful units (or 'elements') such as verses, quotations, cross-references, and other things begin and end.

If you know HTML, you are ahead of the game and already know most of what you need to know to use OSIS; OSIS uses the same pointy-bracket syntax as HTML (or XML to be completely precise). It merely provides a different set of element and attribute names. A few names such as 'p' and 'div' are the same; others are new, such as 'verse.' The core set of elements for OSIS is only slightly larger than the set for HTML 3.2. To be sure, there are some complex cases that we deal with later, but you can do useful work with no more information than is provided in this basic manual.

The second question that is most often asked is: 'Do I need an XML editor to do OSIS?' This question often comes up after a friend of a friend has recommended some editor, and you then checked its price. XML editors vary from free to over $10,000.00 (US), and their ease of use varies greatly.

The basic answer is no, you do not need any special software. You can use any text editor you like to create OSIS documents. Many will even color the tags for you, because they know how to color HTML tags and the languages are similar enough. However, you should have a way to check your documents for errors -- if your editor doesn't know enough about XML to warn you if you misspell a tag, or forget to end some element that you started, you will want to check for errors periodically using an "XML validator".

Many XML editors offer validation of XML documents against a schema. If you don't use an XML editor, you can still validate your OSIS document using a web based validator. These validators allow you to select a file on your local computer and its schema and the service will validate it for your. Two such services are: http://apps.gotdotnet.com/xmltools/xsdvalidator/, and http://tools.decisionsoft.com/schemaValidate.html.

An OSIS-aware text editor will do this checking for you, either on demand or continuously. A better OSIS-aware text editor will provide help by showing you just which elements are permitted at any given place. The best editors also give you the option to see and edit a fully-formatted view on demand, rather than staring directly at pointy-brackets. The choice between the many tools is a personal one, dictated by your working style, level of technical sophistication, goals, budget, and other factors.

Note on examples: This manual offers a number of examples of OSIS markup. The formatting of those examples (indenting, line breaks, etc.) is for ease of reading only. XML processors don't care if all the elements were run one after the other, although most people would find that difficult to read. Liberal use of whitespace and line breaks will make your encoding easier reading for human proofers. That includes you, so use it to your advantage!

For future reference: From time to time you will see paragraphs that start: **For future reference:**. On your first (or even second) reading of this manual you can safely skip those paragraphs. They contain information that you may need later as you become more experienced in using XML or for solving a particular problem.

3. Some authoring tools

The OSIS team is working to adapt free authoring tools that will hide most, if not all, of the markup from the casual user of OSIS. In the meantime, the best way to learn OSIS is to use a simple text editor, such as WordPad or Kedit on Windows, BBEdit or Alpha on MacOS, or vi or emacs on Linux. You can even use a word processor, though any formatting that you do in it won't matter (you would simply save the file as "text only").

Editing is much easier with an editing program that is aware of XML rules in general, and OSIS in particular. For example, rather than seeing literal tags with pointy-brackets, you can have a choice of seeing that, or structural views of your document (say, as a tree or expandable outline), or fully-formatted views to facilitate print layout.

Many products are available that can help you edit XML documents. One style shows the literal XML source file, but colors tags, attributes, and other things to make them stand out. Most such programs also read an XML schema and ensure that you only insert elements and attributes permitted by the OSIS schema (schemas, such as the OSIS one, declare what elements and attributes are permitted where in documents of a particular kind). One free and helpful tool of this kind is jEdit, http://www.jedit.org/ , which runs on most platforms. It can be set up to know about many kinds of files, including XML files, and OSIS in particular.

With such an editor, you can see or print a basic a formatted view by using most any Web browser. Later in this manual are instructions for setting up an OSIS file with a style sheet (generally in CSS) so that typical browsers can render the OSIS text. An OSIS file can also be rendered to HTML or PDF, for example, by using XSLT stylesheets.

There are also more word-processor-like XML editors, which primarily show a formatted view defined by some style sheet. These are mainly commercial. Commercial editors include: XML Spy (http://www.xmlspy.com/, note there is a free version, XML Spy Home, that will work for editing OSIS documents); XMetal (http://www.corel.com), and Serna Editor (http://www.syntext.com). Free editors and other XML tools can be found at Free XML Tools (http://www.garshol.priv.no/download/xmltools/), as well as other XML sites on the WWW.

For high-end layout and typesetting from XML source files, usually a stylesheet language called XSL-FO is used. Popular commercial XSL-FO solutions include: 3b2 (see http://www.3b2.com/), Antenna House (see http://www.antennahouse.com/), and, RenderX (see http://www.renderx.net/. A free XSL-FO processor, FOP (Formatting Objects Processor) is available from the Apache Foundation, (http://xml.apache.org/fop/. Non-XML-based composition systems such as Quark™ and TeX have ways to import XML, but using them for XML composition requires substantial expertise and effort.

4. Your First OSIS Document

Like HTML documents, an OSIS document starts with a header, and is followed by the actual text content. The header identifies the file as being XML, and that it uses the OSIS schema. It also provides places to declare a bibliographical description of the text and of any other works cited; and a place to record a history of editing changes.

Do not be afraid of the following short, but valid, OSIS document! While this encoding may look like a bear, think of it as a Koala bear. Most of what you see is boilerplate that you will simply copy-n-paste once when you start working with a Bible text using OSIS. An OSIS document is simply an XML document that has OSIS specific boilerplate at the top and follows the rules for OSIS markup.

```
<?xml version="1.0" encoding="UTF-8"?>
<osis xmlns="http://www.bibletechnologies.net/2003/OSIS/namespace"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.bibletechnologies.net/2003/OSIS/namespace
            http://www.bibltechnolgologies.net/OSIS/osisCore.2.1.xsd">
<osisText osisIDWork="CEV" osisRefWork="bible" xml:lang="en" canonical="true">
        <header>
            <work osisWork="CEV">
                <title>Contemporary English Version</title>
                <type type="OSIS">Bible</type>
                <identifier type="OSIS">Bible.en.ABS.CEV.1995</identifier>
                <rights type="x-copyright">Copyright 1995
                                    American Bible Society</rights>
                <scope>Esth.1.1-Esth.1.4</scope>
                <refSystem>Bible</refSystem>
            </work>
            <work osisWork="bible">
                <type type="OSIS">Bible</type>
                <refSystem>Bible</refSystem>
            </work>
        </header>
  <div type="section" scope="Esth.1.1-Esth.1.4">
        <title>Queen Vashti Disobeys King Xerxes</title>
          <p>
<verse sID="Esth.1.1-Esth.1.2" osisID="Esth.1.1 Esth.1.2" n="1-2"/>
  King Xerxes of Persia lived in his capital city of Susa and ruled one
  hundred twenty-seven provinces from India to Ethiopia.
<verse eID="Esth.1.1-Esth.1.2"/>
<verse sID="Esth.1.3" osisID="Esth.1.3"/>
  During the third year of his rule, Xerxes gave a big dinner for all
  his officials and officers. The governors and leaders of the provinces
  were also invited, and even the commanders of the Persian and Median
  armies came.
<verse eID="Esth.1.3"/>
<verse sID="Esth.1.4" osisID="Esth.1.4"/>
  For one hundred eighty days he showed off his wealth and spent a lot
  of money to impress his guests with the greatness of his kingdom.
<verse eID="Esth.1.4"/>
          </p>
  </div>
</osisText>
</osis>
```

5. XML and OSIS declarations

In order to be useful (actually to work at all), all text being encoded in OSIS need to be identified as such. This is accomplished by using declarations. Encoding declarations tell the parser (part of the software that reads the file) what type of document it's reading.

The first line:

```
<?xml version="1.0" encoding="UTF-8"?>
```

has three parts:

- **xml**
- **version="1.0"**
- **encoding="UTF-8"**

The first two, **xml version="1.0"** identify the document as being XML and its version of XML. The third one, **encoding="UTF-8"** tells the parser (part of the software that reads the file) what character set has been used with this document. Think of a **character set** as the set of characters that can be displayed in the document. Other encoding declarations are possible but this will be the most common one.

The second through third lines are a very long start-tag for the outermost OSIS element, which is called **<osis>**. All elements in an OSIS document must be declared within the OSIS namespace.

A **namespace** is a method defined by the XML standard to declare the origin of the elements in a document. The XML standard allows for elements from different definitions to be combined in a single document. It also allows different methods for defining elements, including having no definition in a schema at all. To manage those situations, all XML based documents must have a namespace declaration.

There are two ways to use the OSIS namespace and other than remembering to pick one of the two following methods, that is **all** you need remember about it to start encoding texts using OSIS 2.1.1.

OSIS Namespace, Method 1: Copy the following lines after <?xml version="1.0" encoding="UTF-8"?>:

```
<osis
xmlns="http://www.bibletechnologies.net/2003/OSIS/namespace"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.bibletechnologies.net/2003/OSIS/namespace
                    http://www.bibltechnolgologies.net/OSIS/osisCore.2.1.1.xsd">
```

OSIS Namespace, Method 2: Copy the following lines after <?xml version="1.0" encoding="UTF-8"?>:

```
 <osis:osis
xmlns:osis="http://www.bibletechnologies.net/2003/OSIS/namespace"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.bibletechnologies.net/2003/OSIS/namespace
http://www.bibltechnolgologies.net/OSIS/osisCore.2.1.1.xsd">
```

Note with the second method, the last closing element must be: </osis:osis>. The first method is simpler but both are legitimate.

For future reference: The first method declares that the OSIS namespace is the default namespace. The second allows use of the OSIS namespace only if you have the prefix "osis" before the element where the namespace will be used. Namespaces are inherited from their parent elements so there is no practical impact on how you would author an OSIS document.

For future reference: Note that the URL to the location of the OSIS schema is given. This means that you must be connected to the WWW in order to validate a document. For local use, you must have a local copy of the schema in the same directory as your document. In that case, http://www.bibltechnolgologies.net/OSIS/osisCore.2.1.xsd" becomes, osisCore.2.1.xsd".

At this point, the OSIS document has begun. OSIS documents can either be a corpus of multiple works, or merely a single text like the KJV Bible. This sample is a single document rather than a collection of documents (which would start with an **<osisCorpus>** element with multiple **<osisText>** elements inside), so the next element opened is **osisText**:

```
<osisText osisIDWork="CEV" osisRefWork="Bible" xml:lang="en" canonical="true">
```

The **osisText** element has more attributes than are shown here, but these are the most important ones for a basic OSIS document. (For a complete listing of all the attributes of **osisText** or other elements, see its listing in Appendix A: OSIS Elements.

Every **osisText** element is **required** to supply an **osisIDWork** attribute and value. The value will be the short name of what is being encoded, in this case the Contemporary English Version, or CEV. The short name is defined in the **work** declaration for the text, described later.

Every **osisText** element specifies what reference or versification scheme any osisRefs within it refer to. This may or may not be the same work. Depending on how finely you distinguish things, there are several major versification traditions, and countless fine-grained variations. For the present, we identify and reserve names for these major traditional reference systems:

- **NRSVA** New Revised Standard Version with Apocrypha
- **NA27** Nestle-Aland, 27th Edition of the Greek New Testament
- **KJV** King James Version or Authorized Version (AV)
- **LXX** Septuagint
- **MT** Masoretic Text. Hebrew tradition varies in several respects, the best known being that it numbers what is given as a title for Psalms in most English translations as verse 1,and the beginning of the psalm in such a translation as verse 2.
- **SamPent** the Samaritan Pentateuch used a quite different numbering system.
- **Synodal** Russian
- **Vugl** Vulgate
- **Loeb** This system is used for most classical literature, though many major works have other systems as well.

OSIS is developing a schema for declaring versification systems formally, and for declaring some systems in terms of others. This will enable programs to map between systems. However, at this time we merely reserve the names above for some systems we know to be substantially different and important.

The **xml:lang** attribute is required on all **osisText** elements. While the **language** element in **osisWork** allows a wide range of language classification systems to be documented, be aware that **xml:lang** is much more restrictive. The **xml:lang** attribute will only recognize a value from one of the two following lists:

17

- ISO 639: http://lcweb.loc.gov/standards/iso639-2/langhome.html, or
- IANA: ftp://ftp.isi.edu/in-notes/iana/assignments/languages/tags

These codes identify the language of the document. ISO 639 is a standard from the International Organization for Standardization (ISO) and To use anything else with this attribute, you must first prepend a "x-" to the value. To use the SIL Ethnologue code for English, 'eng', for example, it would be written 'x-eng'.

The **canonical** attribute is available on all elements. It has a 'default' value so it does not have to be entered by the encoder if the default value is acceptable. Here it is shown with its default value of **true** on **osisText** so the reader can see it in operation.

When **canonical="true"**, it means that the content of that element is a part of the text being encoded. For example, the "text" of the Bible includes the content of books, chapters, and verses but does not include notes, section-headings added by editors or translators, etc. Therefore, the default value for elements such as **note** is false, as the content of that element has been added by an editor or author to the text being encoded. *It should be explicitly noted that the value of the* **canonical** *attribute should not be used to reflect theological judgment about the content of a text, but merely to distinguish between what has been added to the text and what has not.*

In most cases use of the **canonical** attribute is straightforward, and the default values will almost always produce the intended result. However, there will arise truly difficult cases: for example, one may be encoding an ancient text with annotations of its own. In that case those notes would be canonical, while any added by the current editor would not be. In such cases, the practice chosen and its rationale should be described in the work's documentation.

A common case where users need to watch the value of the canonical attribute is the use of **<title>** in the Psalms. The default value for canonical on **<title>** is false, but in the case of individual Psalms, that value should be true. The default was set to false as most of the titles for sections and the like that are found in biblical texts are in addition to the actual text.

The **canonical** attribute is available on all elements. The value of this attribute is "inherited," that is once it is set, any subelement of that element inherits the same setting. For example:

```
<osisText osisIDWork="CEV" osisRefWork="Bible" canonical="true"
xml:lang="en">
<header>...header elements omitted...</header>
 <div type="section" scope="Esth.1.1-Esth.1.4">
    <title canonical="false">Queen Vashti Disobeys King Xerxes</title>
      <p>
<verse sID="Esth.1.1-Esth.1.2" osisID="Esth.1.1 Esth.1.2" n="1-2"/>
 King Xerxes of Persia lived in his capital city of Susa and ruled one
 hundred twenty-seven provinces from India to Ethiopia.
<verse eID="Esth.1.1-Esth.1.2"/>
```

The **canonical** attribute on the **title** element is set to false because the title in question does not appear in the Bible text. If the **title** element contained other elements, they too would be considered as having a **canonical** attribute setting of false.

The next element, **p** element is a child of the **osisText** element and so it inherits the 'true' Because of this inheritance, encoders will seldom need to make this attribute explicit. In **osisText** this attribute is set to a default value of "true", while in header, note, and reference that setting is overridden by setting the value of

that attribute to be "false."

In books other than the Bible, a similar distinction holds: the text proper of Herodotus' Histories must be contained in elements with the **canonical** attribute set to 'true', while notes, header data, and the like must have the same attribute set to 'false'.

In most cases use of the **canonical** attribute is straightforward, and we expect that the default values will almost always produce the intended result. However, there will arise truly difficult cases: for example, one may be encoding an ancient text with annotations of its own. In that case those notes would be canonical, while any added by the current editor would not be. In such cases, the practice chosen and its rationale should be described in the work's documentation in the **header** element.

5.1. Elements allowed in osisText elements

The **osisText** element contains the following elements,:

• **header** • **titlePage** (optional) • **div**

6. The OSIS header

The first element following **osisText** is **required** to be **header**. The **header** contains the **revisionDesc**, **work**, and **workPrefix** elements for a particular work. These elements must be entered in that order, although each may occur an unlimited number of times.

In other words, the order **always** is:

- **<revisionDesc>** elements (any number of them), followed by:
- **<work>** elements (any number of them), followed by:
- **<workPrefix>** elements (any number of them).

6.1. revisionDesc

The **revisionDesc** element is used to record changes or edits to the text and should be used every time significant editing is done. Each **revisionDesc** element must contain a **date** element which says when those edits were completed, in the form

```
yyyy.mm.ddThh.mm.ss
```

Note that all fields must have exactly the number of digits shown (4-digit year, 2-digit month, etc.). It is permissible to omit the time and the preceding "T", thus giving just a date. For example, December 25th of 1999 CE would be:

```
1999.12.25
```

A **date** element in the revision description is followed by any number of **p** (paragraph) elements, in which the changes made are summarized. The person responsible for making the changes should also be identified, using the **resp** attribute on the **revisionDesc** element. The **resp** attribute records who made a change or edit to the text.

Recommended practice is that more recent **revisionDesc** elements appear earlier in the document. That is, entries should occur in reverse chronological order. For example:

```
    <revisionDesc resp="sjd"><date>2003.09.11</date>
    <p>Filling in the gaps. Adding some info for 2.0 as defined
at the Calvin College meetings.</p>
    </revisionDesc>
    <revisionDesc resp="sjd"><date>2003.07.01</date>
    <p>sjd: Annotated alpha list of elements. Reworked reference and
work sections and added type, scope, and explanations of type and
subtype for work. Explained more elements and attributes.</p>
    </revisionDesc>
    <revisionDesc resp="sjd"><date>2003.06.17</date>
    <p>Wrote conformance section. Added lists of elements and
attributes, USMARC list. Inserted placeholders for doc on all element
types. Got document back to XML WF. Wrote CSS stylesheet.</p>
    </revisionDesc>
<revisionDesc resp="pld"><date>2005.11.01</date>
    <p>Conformed references to new OSIS 2.1 schema.</p>
    <p>Added new examples to the text.</p>
```

```
        <p>Added mapping to USFM codes</p>
</revisionDesc>
```

Note that a separate **revisionDesc** element is used for each person responsible for changes in the text. Here, 'sjd' and 'pld' refer to Steve DeRose and Patrick Durusau, respectively. The use of such abbreviations is common but by no means required.

Within a **revisionDesc** element, following the **date** element, it is possible to have as many **p** elements as are necessary to describe the revision. It is better to have too full of a description of a revision than to have too little information about a revision. Users should always err on the side of too many comments. Revision comments will not appear in a normal rendition of the text.

6.2. work

A **work** element provides information comparable to that found on the title page of a printed work, using the fields defined by the Dublin Core Initiative (see http://dublincore.org/).

The **work** element in the **header** with an **osisWork** attribute that matches the **osisIDWork** in the **osisText** element identifies the work in which it occurs, like the title page in a printed work. For example:

```
<osisText osisIDWork="CEV" osisRefWork="Bible" xml:lang="en" canonical="true">
      <header>
            <work osisWork="CEV">
```

Note that the match between **osisIDWork="CEV"** in **osisText** and **osisWork="CEV"** in the **work** element links this **osisText** to this particular **work** element.

Other **work** elements (ones that follow the first one) identify other works, much like a citation in a footnote or bibliography in a printed work. Each assigns a *local name* to that work, using the **osisWork** attribute. Works so declared can then be referred to in **osisID**s or **osisRef**s throughout the text. For Bibles, this should generally be the accepted acronym or abbreviated form of the translation's name (some standard version abbreviations are listed above). No periods, hyphens, spaces, or colons are allowed in short names.

6.3. osisWork

The **osisWork** attribute of each **work** element provides a short name used to refer that work and its declaration as a whole. As of version 2.1, OSIS specifies the recommended format described below for those short names. When using the short name from an OSIS work element within the Dublin Core **identifier** element, the **type** attribute must be set to "OSIS".

The first **work** element in the **header must** identify the work that is being encoded. It is used as the default value for all **osisIDs** and **osisRefs** that do not have values or default values declared otherwise.

6.4. Work Prefix Defaults

Each work defined in the header is required to provide a short name for itself in the **osisWork** attribute. These can then be used as prefixes (similar to XML namespace prefixes) on **osisID**, **osisRef**, and other attributes throughout the rest of the document.

21

In OSIS versions through 2.0, specific attributes were provided to set a default work prefix for **osisID**s (**osisIDWork** on the **osisText** element) and for **osisRef**s (**osisRefWork** on the **osisText** element). These attributes remain available in OSIS 2.1, but a more general defaulting mechanism has been added.

In OSIS version 2.1 and later, the **workPrefix** element was added to make it possible to specify a default work prefix for the attributes on any element in an OSIS document. The **workPrefix** element appears at the end of the header, after all the **work** elements. Any number of **workPrefix** elements are allowed, each of which sets the default work prefix for a particular attribute on a particular element type. For example:

```
<workPrefix path="//note/@annotateRef" osisWork="Bible.KJV"/>
```

This declaration indicates that the default work prefix on all **annotateRef** attributes of **note** elements is to be "Bible.KJV". No colon is to be included (the colon is used to separate a work prefix from the rest of a reference when the work prefix is explicit rather than defaulted).

The syntax of the **path** attribute is taken directly from the W3C XPath Recommendation, and can be correctly interpreted by any conforming XPath processor. (A **path** attribute is a means to specify a an element in an XML document. The leading "//" means that this will find any **note** element in the document, wherever it may be found.) However, the form shown above is the *only* form permitted at this time; more complex XPaths are not permitted. In other words, the **path** attribute must consist of "//", an element type name, "/@", and an attribute name. If a more detailed defaulting mechanism is required in the future, it will likely be provided by permitting a wider range of XPath's features.

The **w** element has attributes which record the morphology and base form of a word (the **morph** and **lemma** attributes respectively). Because the **w** element is so frequent when used, defaulting the prefix (which points to a work defining the morphological of lexical system used) can save a lot of data entry effort, which will minimize typing errors.

7. work

7.1. Introduction

The elements that can appear in the **work** element are organized in the OSIS schema beginning with those found in Dublin Core metadata followed by ones that are specific to OSIS. Dublin Core metadata provides basic information about a **work** and the OSIS elements are OSIS specific additions to that basic information.

Each **work** element describes a single publication using several pieces of information, primarily **title**, **creator**, **date**, **publisher**, **identifier** and **language**. There are additional standard "Dublin Core" fields that can be used, plus a few OSIS-specific additions (further information on the Dublin Core system may be found at http://www.dublincore.org). All of the elements in **work** may be repeated as necessary, but must be encoded in the order shown here. For example:

```
<work osisWork="thisWork">
    <title>New International Version</title>
    <type type="OSIS">Bible</type>
    <identifier type="OSIS">Bible.en.NIV.1984</identifier>
    <rights type="x-copyright">American Bible Society 1994</rights>
    <refSystem>Bible</refSystem>
</work>

<work osisWork="EG">
    <title>Egyptian Grammar</title>
    <creator role="aut">Alan Gardiner</creator>
    <contributor role="dte">Francis Llewellyn Griffith</contributor>
    <date event="original" type="gregorian">1927</date>
    <date event="eversion" type="gregorian">2003</date>
    <type type="x-grammar">Grammar</type>
    <publisher>Griffith Institute, Ashmolean Museum, Oxford</publisher>
    <language type="ISO-639">EN</language>
    <language type="Ethnologue">EG-ancient</language>
    <identifier type="ISBN">0900416351</identifier>
    <identifier type="LCCN">95230980</identifer>
</work>

<work osisWork="CPV">
    <title>Cotton Patch Version of Luke and Acts: Jesus' Doings and
the Happenings</title>
    <creator role="aut">Clarence Jordan</creator>
    <date event="original" type="gregorian">1969</date>
    <date event="eversion" type="gregorian">2003</date>
    <type type="x-bible">Bible</type>
    <publisher>Association Press
       <name type="place">New York, NY</name></publisher>
    <language type="ISO-639">EN</language>
    <identifier type="ISBN">0809617250</identifier>
    <identifier type="LCCN">69-18840</identifer>
    <scope osisRef="Luke Acts" />
</work>
```

7.2. Dublin Core Elements

7.2.1. title

A **title** element *must* be provided in the **work** element and contain the main title of the work. Additional titles may also be specified, using the **type** attribute to identify them as main, sub, part, or another kind of title.

Caution: The xml:lang attribute on **<osisText>** controls the language of all later elements unless a new **xml:lang** attribute value is set. In other words, if there are multiple titles in different languages, then each **<title>** element should have the appropriate **xml:lang** attribute value. When used for more than simply different **<title>** elements, that is to distinguish other elements, the **xml:lang** attribute can serve as the basis for maintaining multiple language versions of a single document. One example would be to have the same **note** elements in multiple languages for a single biblical text in English or French and to produce the text by selecting the language of the **note** elements to be included.

7.2.1.1. Title Placement

One of the more difficult and variable problems faced by users who want to render an encoded document is the placement of titles in the work. Since titles occur for the work as a whole, as well as elsewhere, the following values have been defined for the **placement** attribute on the **title** element:

- leftHead
- centerHead
- rightHead
- insideHead
- outsideHead
- leftFoot
- centerFoot
- rightFoot
- insideFoot
- outsideFoot

7.2.2. creator

The **creator** element is used to specify the person(s) or organization(s) who are primarily responsible for the intellectual content of a work. The **role** attribute must specify the role the primary responsible party played. The most common values would be aut (author), edt (editor), cmm (commentator), trl (translator). A short list of such codes appears along with the complete set appears in the appendices under *USMARC Relator Codes*.

7.2.3. contributor

Many people may contribute to a work in roles other than the primary roles listed under creator. They should be encoded using the **contributor** element. Their specific role should be recorded in the **role** attribute of their **contributor** element.

There is an optional attribute **file-as** on this attribute. Contributors may have names or other identifiers that are used to find texts with particular contributors.

24

For example, if encoding a Bible that has illustrations by Gustave Doré, the following **contributor** element would be appropriate:

```
<contributor role="ill">Gustave Dor&#xE9;</contributor>
```

The **role** attribute has the value "ill" which stands for "illustrator" as documented in the *USMARC Relator Codes*.

Note: The **é** is a character entity. Character entities are a way to enter characters that are not available on a local keyboard. It displays after processing as the character it represents. In this case, an "e" with an acute accent.

7.2.4. date

Date elements in the **work** element record significant dates in the production or publication process. Use the **event** attribute to identify the particular date contained in each **date** element. The defined events are:

- **edition** The date of publication of the referenced or source edition
- **eversion** The revision date of the present electronic edition
- **imprint** The printing date of the referenced or source edition
- **original** The original publication date of the first edition

The **type** attribute is used to identify the calendrical system in which the date is expressed, from the list: **Chinese**, **Gregorian**, **Islamic**, **ISO**, **Jewish**, and **Julian**. At this time, OSIS only defines a syntax for ISO dates: yyyy.mm.dd.

For example:

```
<date event="edition" type="ISO">1749</date>
<date event="eversion" type="ISO">2003.12.30</date>
```

For processing purposes, OSIS compliant software will assume ISO dates unless the **type** attribute indicates otherwise. See *Date Formats* for further details on the encoding of date information.

7.2.5. publisher

The **publisher** element in the **work** element is used to identify the publisher of a particular work. If a work is co-published (at the same time) by more than one publisher, then multiple **publisher** elements should be used.

If a publication history of a work is to be encoded, then multiple **work** elements must be used. An example would be the publication of *Brown, Driver, Briggs*, better known to Hebrew students as *BDB*, which exists in several editions. Each publication event constitutes a separate work.

```
<work osisWork="bdb1907">
...
<publisher>Houghton Mifflin</publisher>
<date>1907</date>
...
</work>
<work osisWork="bdb1955">
```

```
...
<publisher>Clarendon Press</publisher>
<date>1955</date>
...
</work>
```

While each edition of this work retained the same title, each one introduced corrections and changes that were not present in prior editions.

7.2.6. language

A **language** element may be provided for each language used substantially in a work. The language may be specified using an ISO 639 or ISO 639-2, or SIL Ethnologue code. The **type** attribute must be set to **IANA**, **IETF**, **ISO-639-1**, **ISO-639-2**, **ISO-639-2-B**, **ISO-639-2-T**, **LINGUIST**, or **SIL**. In the rare case that none of these is sufficient, a prose description should be inserted in the element and the **type** attribute set to **other**.

The use of a language in a text can be specifed using the **use** attribute of the **language** element. The permitted values are:

- base
- didactic
- interlinear
- original
- quotation
- source
- target
- translation

7.2.7. type

The **type** element encodes he nature or genre of the content of the resource. This element includes terms describing general categories, functions, genres, or aggregation levels for content. Dublin Core's recommended best practice is to select a value from a controlled vocabulary (for example, the DCMI Type Vocabulary, see http://dublincore.org/documents/dcmi-type-vocabulary/). OSIS does not provide such a controlled vocabulary at this time.

If you encode this element, the controlled vocabulary in use should be identified via the **type** attribute (for example, **<type type="x-DCMI">**).

The **type** attribute should **not** be used to describe a resource as a 'book', 'electronic file', or other form of the work. Use the **format** element instead.

For example, assume that you are encoding the first English language New Testament printed in the United States (by Robert Aitken) from a microfilm copy, the **type** and **format** elements could be used as follows:

```
<type type="x-DCMI">text</type>
<format>microfilm</format>
```

7.2.8. identifier

The **identifier** element provides one formal identifier for the work. The values to be entered for the **type** attribute on the **identifier** element are shown in bold. Note that these values must be entered exactly as shown. XML is case sensitive, that is to say, **DEWEY** is **NOT** equal to **Dewey**. Enter **Dewey** and you will get an error message.

A work, represented by a **work** element, can have more than one **identifier** element. That is to say it may have different identifiers in different identification systems. For example, a book could have recorded in its **work** element, as having two **identifier** elements, one for its **ISBN**, <identifier type="ISBN">0-310-92955-5</identifier>, and one for its Library of Congress Control Number, <identifier type="LCCN">2002107776</identifier>.

- **DEWEY** Dewey Decimal System
- **DOI** Digital Object Identifier
- **ISBN** International Standard Book Number
- **ISSN** International Standard Serial Number
- **LCC** Library of Congress Classification
- **LCCN** Library of Congress Control Number (also known as "Library of Congress Card Number")
- **OSIS** Open Scriptural Information Standard
- **SICI** Serial Item and Contribution Identifier
- **URI** Uniform Resource Identifier
- **URL** Uniform Resource Locator
- **URN** Uniform Resource Name

ISBN and LCCN numbers must be recorded without spaces or hyphens. ISBNs must contain ten digits (that is, they must include the final check digit).

If this is an encoding of a specific published work, we strongly recommend the ISBN be used to identify the work. This number must, if available, be specified in the **identifier** element for the work.

The following examples show **identifier** elements used along with their **type** attribute to provide an identifier for a work, in this case, the "Cotton Patch Version of Luke and Acts" noted above:

```
<identifier type="ISBN">0809617250</identifier>
<identifier type="LCCN">69-18840</identifier>
```

Without the proper **type** attribute, a reader or computer only has a string of numbers, which could be from almost any system of identifiers. The **type** attribute plays an important role in making sure the information is understandable by yourself or others.

7.2.9. coverage

This element may be used to specify the spatial location (a place name or geographic coordinates), temporal period (a period label, date, or date range) or jurisdiction (such as a named administrative entity) to which the work applies. For example, an edition of Herodotus could be specified as Greek/Hellenic, Classical Period. Or a study of Medieval Bibles could declare coverage as Medieval.

```
<coverage>Greek/Hellenic, Classical Period</coverage>
<coverage>Medieval</coverage>
```

7.2.10. description

An account of the content of the resource.

Examples of **description** include, but are not limited to: an abstract, table of contents, reference to a graphical representation of content or a free-text account of the content.

```
<decription>This encoding of the first New Testament printed in the
colonies of the United States follows the spelling conventions of that
period.</description>
```

Information in the **description** element cannot be automatically processed for details about the text. It's use should be limited to providing a reader friendly description of the text.

The only defined value for the **type** attribute of the **description** element is **usfm**. If other values are required, they must be preceded by "x-," thus, **x-pld**.

7.2.11. format

The physical or digital manifestation of the resource.

Typically, **format** may include the media-type or dimensions of the resource. Format may be used to identify the software, hardware, or other equipment needed to display or operate the resource. Examples of dimensions include size and duration. Recommended best practice is to select a value from a controlled vocabulary (for example, the list of Internet Media Types (MIME) http://www.iana.org/assignments/media-types/ defining computer media formats).

OSIS documents will typically have the following **format** element with the values as shown:

```
<format type="x-MIME">text/xml</format>
```

The **type** attribute with the value **MIME** means that the content of the **format** element has been taken from the list of Internet Media Types mentioned above. The "x-" preceding the attribute value is required. This element has no OSIS predefined values for its **type** attribute.

7.2.12. relation

A reference to a related resource.

Recommended best practice is to identify the referenced resource by means of a string or number conforming to a formal identification system.

For example, an OSIS document that is encoding only a portion of Spurgeons Sermons could indicate that it is part of a larger set of sermons as follows:

```
<relation type="x-isAPartOf">SpurgeonsSermons</relation>
```

The "x-" preceding the attribute value is required. This element has no OSIS predefined values for its **type** attribute.

7.2.13. rights

Information about rights held in and over the resource.

Typically, the **rights** element will contain a rights management statement for the resource, or reference a service providing such information. Rights information often encompasses Intellectual Property Rights (IPR), Copyright, and other property rights. The content of the **rights** element is informative only. Legal rights and penalties for violation of those rights vary from jurisdiction to jurisdiction. Reuse of any resource should be done only after obtaining the necessary rights and permissions or ascertaining that none is required.

```
<rights type="x-copyright">Copyright 1995 American Bible Society</rights>
```

The "x-" preceding the attribute value is required. This element has no OSIS predefined values for its **type** attribute.

7.2.14. subject

A topic of the content of the resource.

Typically, the **subject** will contain keywords, key phrases or classification codes that describe a topic of the resource. Recommended best practice is to select a value from a controlled vocabulary or formal classification scheme.

7.2.14.1. subject classification systems

The **type** attribute on **subject** allows the user to specify the classification system the subject entered can be found.

```
<subject type="ATLA">Fathers of the Church</subject>
```

The above means that the subject "Fathers of the Church" is a subject found in the listing of subjects maintained by the American Theological Libraries Association (ATLA). To assist users, an admittedly partial list of the more well known subject classification systems has been prepared by the OSIS project. Those systems with their abbreviations for use with an OSIS encoded text are as follows:

- **ATLA** American Theological Libraries Association
- **BILDI** Bibelwissenschaftliche Literaturdokumentation Innsbruck
- **DBC** Dutch Basic Classification
- **DDC** Dewey Decimal Classification
- **EUT** Estonian Universal Thesaurus
- **FGT** Finnish General Thesaurus
- **LCC** Library of Congress Classification
- **LCSH** Library of Congress Subject Heading
- **MeSH** Medical Subject Headings
- **NLSH** National Library Subject Headings (National Library of Poland)
- **RSWK** Regeln für den Schlagwortkatalog
- **SEARS** Sears List of Subject Headings
- **SOG** Soggettario
- **SWD_RSWK** Swiss National Library

- **UDC** Universal Decimal Classification
- **VAT** Vatican Library

For classification systems not listed, insert the classification system with a leading "x-" in the **type** attribute, and notify the OSIS editors (osis-editors@bibletechnologieswg.org) so that system can considered for addtion to a future revision of the schema.

7.2.14.2. source

A reference to a resource from which the present resource is derived.

The present resource may be derived from the resource identified by the **source** element in whole or in part. Recommended best practice is to identify the referenced resource by means of a string or number conforming to a formal identification system.

For example, assume you are encoding the New Testament as printed by Robert Aitken (1782) and wish to indicate the source you used for that text:

```
<source>Emory University Library: MICFICHE 1367 NO. 17101 </source>
```

7.2.15. Non-Dublin Core Elements and Attributes in the Work Declaration

7.2.15.1. scope

The content of **scope** element(s) must be in the form of an **osisRef** value, and that defines what part of the titled work occurs in this electronic edition. For example, an edition may consist of only the New Testament and Psalms, or of only a single book. Contiguous ranges may be specified using the hyphen notation described later for **osisRef**s in general; discontiguous ranges are specified by including multiple values using **osisRef** syntax. These should be in canonical order, but that is not required.

7.2.15.2. castList

The **castList** element is composed of any number of **castGroup** elements.

7.2.15.3. castGroup

The **castGroup** element provides an easy way to consolidate information about the characters in a dramatic text for later reference in the encoding of that text. The **castGroup** element is composed of **head** and **castItem** elements. See the material under *Dramatic Texts* for details and examples on using the **castList** and related elements.

7.2.15.4. teiHeader

The content of a **teiHeader** element is not processed as part of the OSIS schema. It is provided for cases where a TEI header file or portion of a text needs to accompany a text that is now being encoded in OSIS markup.

There are few Bibles that have been encoded using TEI but the following is the **teiHeader** from one held by the Oxford Text Archive:

```
<teiHeader type="ISBD(ER)">
<fileDesc>
    <titleStmt>
        <title type="uniform">Bible. English. Authorized</title>
        <title type="main">
        The Bible [Electronic resource] : King James version
        </title>
     <respStmt>
        <resp>creation of machine-readable version</resp>
        <name>Project Gutenberg</name>
     </respStmt>
     </titleStmt>
     <extent>
        <seg type="designation">Text data</seg>
        <seg type="size">(1 file : ca. 4.72 megabytes)</seg>
     </extent>
     <publicationStmt>
        <distributor>
        <name key="ota" type="organisation">Oxford Text Archive</name>
        <name type="place">Oxford</name>
        <address>
        <addrLine>
        <name key="oucs" type="organisation">Oxford University Computing
Services</name>
        </addrLine>
        <addrLine>13 Banbury Road</addrLine>
        <addrLine>Oxford</addrLine>
        <addrLine>OX2 6NN</addrLine>
        <addrLine>
        <name type="email">info@ota.ahds.ac.uk</name>
        </addrLine>
        </address>
        </distributor>
        <idno type="ota">bibl1691</idno>
        <availability status="free">
        <p>
        Freely available for non-commercial use provided that this header is
        included in itsentirety with any copy distributed
        </p>
        </availability>
        <date>1994-02-13</date>
     </publicationStmt>
     <notesStmt>
        <note>Mode of access: Online. OTA website</note>
        <note>
        Title proper taken from title page of electronic text
        </note>
        <note>
        This text was formerly the second Project Gutenberg version of
        their 10th edition of the King James Bible
        </note>
     </notesStmt>
     <sourceDesc>
        <bibl>Not recorded</bibl>
     </sourceDesc>
</fileDesc>
<encodingDesc>
   <projectDesc>
```

```
   <p>
     Project Gutenberg produces electronic texts that they hope extremely
     large portions of the audience will want and use frequently. In the
     same vein, Project Gutenberg has avoided requests, demands, and
     pressures to create ?authoritative editions.? Their goal is to
     release electronic texts that are 99.9% accurate in the eyes of the
     reader in general, rather than the scholar in particular
   </p>
 </projectDesc>
 <samplingDecl>
   <p>
   </p>
 </samplingDecl>
 <editorialDecl>
   <p>Encoding format: SGML TEI Lite</p>
   <p>
   All direct speech has been represented by quotation entity references
   </p>
 </editorialDecl>
 <tagsDecl>
   <tagUsage gi="body" occurs="1"/>
   <tagUsage gi="div1" occurs="2"/>
   <tagUsage gi="div2" occurs="66"/>
   <tagUsage gi="div3" occurs="1189"/>
   <tagUsage gi="docTitle" occurs="1"/>
   <tagUsage gi="front" occurs="1"/>
   <tagUsage gi="head" occurs="66"/>
   <tagUsage gi="p" occurs="31102"/>
   <tagUsage gi="text" occurs="1"/>
   <tagUsage gi="titlePage" occurs="1"/>
   <tagUsage gi="titlePart" occurs="2"/>
 </tagsDecl>
 <refsDecl>
    <p>Testaments (div1) bear IDs in the form O and N</p>
    <p>Books (div2) bear IDs in the form O1 and N1</p>
    <p>Chapters (div3) bear IDs in the form O1.1 and N1.1</p>
    <p>Verses (p) bear Ns in the form O1.1.1 and N1.1.1</p>
 </refsDecl>
 <classDecl>
    <taxonomy id="OTASH">
       <bibl>Oxford Text Archive Subject Headings</bibl>
    </taxonomy>
    <taxonomy id="LCSH">
       <bibl>Library of Congress Subject Headings</bibl>
    </taxonomy>
 </classDecl>
</encodingDesc>
<profileDesc>
  <creation>
     <date>1992</date>
  </creation>
  <langUsage>
     <language id="eng">English</language>
  </langUsage>
  <textClass>
     <keywords scheme="OTASH">
        <term type="genre">Sacred texts</term>
     </keywords>
     <keywords scheme="LCSH">
        <term type="genre">
        </term>
     </keywords>
```

```
            </textClass>
      </profileDesc>
      <revisionDesc>
         <change>
            <date>2000-10-12</date>
            <respStmt>
               <resp>Cataloguer</resp>
                  <name>Webb, Anton</name>
            </respStmt>
               <item>
                  Header changed in accordance with ISBD(ER) guidelines and expanded.
                  Validated in the TEI Lite DTD using XMetal 2.0
               </item>
         </change>
         <change>
            <date>?</date>
            <respStmt>
               <resp>?</resp>
               <name>?</name>
            </respStmt>
                <item>Text converted to SGML TEI Lite DTD</item>
         </change>
         <change>
            <date>1994-02</date>
            <respStmt>
               <resp>edt (Editor)</resp>
                 <name>Triggs, Jeffery</name>
            </respStmt>
                <item>
                 Preliminary tagging of a plain ASCII text in the SGML OTA DTD
                </item>
         </change>
      </revisionDesc>
   </teiHeader>
```

For further information on the TEI in general or meaning and usage of the elements in a TEI header, please see: http://www.tei-c.org.

7.2.15.5. refSystem

Since versification systems differ between Bible translations (the original texts had no versification at all), it is important to note here which versification system is being followed for a particular edition of the Bible that is being encoded. Works of classical authors as well may have differing reference systems.

The value here should be the **osisWork** attribute value from a work declared in a **work** element in the **header** of the document.

7.2.16. Elements allowed in work elements

The **work** element contains the following elements, which must appear in this order, although they can each be repeated as many times as necessary:

• **title** • **contributor** • **creator** • **subject** • **date** • **description** • **publisher** • **type** • **format** • **identifier** • **source** • **language** • **relation** • **coverage** • **rights** • **scope** • **castList** • **teiHeader** • **refSystem**

7.3. Date Formats

All dates in the header and in attributes should be in format defined below. It uses a period as the field separator, and has features for dates BCE, approximate dates, date ranges, recurring dates (as used in many daily devotionals), weekly dates, and named times of day (used in prayer books). There are three standard date formats; the prefixes that identify them are reserved, and may not be redefined:

- yearly.yyyy.mm.ddThh.mm.ss

  Any number of fields may be left off from the right end; for example, if the seconds are dropped (along with the preceding period), the time refers to the entire minute specified; if the entire time section is left off (along with the preceding "T"), the string refers to the entire day.

  The year must always have four digits. However, the year may be entirely omitted to indicate dates that apply to any year, such as in a book of 365 daily readings.

  To indicates years before the common era, add an underscore ("_") before the first digit of the year (immediately following **yearly.**).

  The entire date/time string (possibly including a leading underscore) may be preceded by "~", indicating that the time is approximate. No means is provided to express just how approximate a time may be.

- weekly.n

  When readings or other materials are specified as being for particular days of the week, this format must be used. The 'n' value may range from 1 to 7; 1 indicates Monday, in accordance with ISO 8601:2000.

- monthly.n

  To organize readings or other materials for specific months, the following format must be used. The 'n' value may range from 1 to 12; where 1 indicates January.

- yearly.n

  To organize readings or other materials for a multi-year reading cycle, the following format must be used. The 'n' value is not constrained but the common lectionary cycle is three years, giving a range of 1 to 3.

As an alternative to quantitative times, a small set of named times is provided, which can be specified in place of the entire (post-"T") time section (the "T" itself remains). For example:

```
yearly.06.04T~(Vespers)
```

would be the identifier for a prayer, reading, or other work to be used at Vespers on June 4 of any year. The named times (which are case-sensitive) include: Vigils, Matins, Lauds, Terce, Sext, None, Vespers, Compline; Sunrise, Sunset; Morning, Afternoon, Evening, Night; AM, PM; Fajr, Zuhr, _Asr, Maghhrib, _Isha, Lail, Dzuha, _Id.

Some works will be primarily organized by dates and times: for example, lectionaries, daily devotionals, prayer books, historical time lines, etc. In such works, use the **osisID** attribute to identify a retrievable portion of that work. The value should the the applicable time in one of the formats just shown.

For example, a book of daily prayers, which is keyed to the calendar date of the year might have each **div** element with an **osisID** written as follows:

```
<div osisID="12.25">daily prayers here in OSIS markup</div>
```

This **osisID** allows retrieval of the prayers for the 25th day of December, without regard to any particular year.

Typically, such works are organized in chronological order of the times specified; however, that is a user or publisher requirement and not one imposed by OSIS.

For future reference: Note that **osisID**s are unconstrained except by their reference systems. Thus, reference systems that rely upon frame references (for multimedia) or more precise timing divisions are available to any user who wishes to declare the necessary reference system.

7.4. Title Pages

OSIS 2.1 introduced the **titlePage** element to enable the encoding of 'static' title pages. It should be noted that the title page, table of contents, for example, of this manual were automatically generated from information in the header and the titles of the various chapters. Automated generation of content is the preferred method for such parts of a work as it avoids extra data entry and inconsistency between various parts of a document. It does, however, require somewhat greater skills at writing stylesheets and may not always meet all needs.

If a 'static' title page is deemed necessary, then the **titlePage** element should be used. This element uses the following elements from the header: **title**, **contributor**, **creator**, **subject**, **date**, **description**, **publisher**, **type**, **format**, **identifier**, **source**, **language**, **relation**, **coverage**, which are explained in the material on the **header** section. Three additional elements are allowed, which are **figure**, **milestone**, and, **p**. Due to the complexity of title pages, all of these elements may occur in any order inside the **titlePage** element.

If the **titlePage** element is used, the elements from the **header** must also appear in the **header** element with the same content as they have in the **header**. XML validation software will not detect this as an error and so it is the responsibility of the author to insure this requirement is meet.

The **titlePage** element can occur within the **osis**, **osisText**, and, **osisCorpus** elements.

Users just starting with OSIS should use a minimum header and a simple **titlePage** element until they have gained some experience with text encoding and determining what is, or perhaps more importantly, what is not useful to have encoded in a work.

A simple title page:

```
<titlePage>
  <title type="main">The Holy Bible
    <title type="sub">Revised Standard Version
       <title type="sub">Containing the Old and New Testaments</title>
```

```
        </title>
    </title>
    <description>Catholic Edition</description>
    <description>Prepared by the Catholic Biblical Association of
                 Great Britian</description>
    <description>With a Foreword by His Eminence John Cardinal Heenan
                 Archbishop of Westminister</description>
    <publisher>Nelson</publisher>
</titlePage>
```

7.4.1. Elements allowed in titlePage elements

The **titlePage** element contains the following elements, which may appear in any order, and may be repeated as many times as necessary:

• **contributor** • **coverage** • **creator** • **date** • **description** • **figure** • **format** • **identifier** • **language** • **milestone** • **p** • **publisher** • **relation** • **source** • **subject** • **title** • **type**

7.5. The div Element

The basic high level divisions of an OSIS document are encoded using the **div** element. The **div** element has a **type** attribute that in a Bible might have the value of **book**, which in a hymnal might have the value **Advent** (to indicate a separate grouping of hymns for the season of Advent). **div** elements can occur within other **div** elements to any number of levels. The first and outermost **div** should occur immediately after the end of the header. For example,

```
<div type="book" osisID="Gen">
    <title type="main" short="Genesis">Genesis</title>
      <chapter osisID="Gen.1" n="1">
          <verse sID="Gen.1.1" osisID="Gen.1.1" n="1"/>In the
          beginning...<verse eID="Gen.1.1"/>
          <verse sID="Gen.1.1" osisID="Gen.1.2" n="2"/>The earth was
          formless and void...<verse eID="Gen.1.2"/>
          ...
      </chapter>
</div>
```

The **div** element is used for many top-level components, and makes heavy use of the **type** attribute. The pre-defined types include the most common major divisions found in present-day Bibles and related works, for example:

• **acknowledgement** • **appendix** • **article** • **back** • **body** • **book** • **bookGroup** • **chapter** • **colophon** • **commentary** • **concordance** • **front** • **glossary** • **index** • **introduction** • **map** • **preface** • **section** • **subSection**

The main body of a Bible will typically consist of **div** elements of **<div type="bookGroup">** (such as each Testament, the Apocrypha, and perhaps smaller groups such as the Pentateuch, the Minor Prophets, etc), plus any front and back matter divisions (the selection of which varies greatly between editions).

The books of the New Testament would be grouped as follows:

```
<div type="bookGroup">
<title>New Testament</title>
...<div>s for individual books here...
</div>
```

Within each **div** of **type="bookGroup"**, there will typically be book division types corresponding to each included Canonical or deutero-canonical book. The **book** division type, **<div type="book">**, may contain **<div type="majorSections">** (such as the sub-books in Psalms, Deuteronomy, Jeremiah), **<div type="sections">** (typically topical divisions with headings), and **<div type="subSections">** (occasional minor divisions within sections). A specific **chapter** element is provided and encouraged in works other than the Bible.

The **head** element follows the **div** element and is used to provide the title for a book or chapter.

Expanding the New Testament example to include the first two Gospels:

```
<div type="bookGroup">
<head>New Testament</head>
<div type="book">
      <head>Matthew</head>
            <chapter>
                  <head>Chapter 1</head>
                  ...text of chapter 1...
            </chapter>
</div>
<div type="book">
      <head>Mark</head>
            <chapter>
                  <head>Chapter 1</head>
                  ...text of chapter 1...
            </chapter>
</div>
</div>
```

It should be noted that **head** elements in the example will be selected by the rendering process to produce both the titles that appear at the begining of each book or chapter, as well as being part of the running header on each page that gives the beginning and ending verse references for that page. Such automated rendering is less error prone than manual entry and allows the output to be shaped according to the delivery needs. For example, running headers on a text delivered to a cellphone would not be useful or perhaps even meaningful.

Below this point typical texts switch from successive levels of **div** elements, to more specific markup such as paragraphs, lists, quotations, inscriptions, and the like. At this level the markup of quotations and paragraphs, for example, begins to interact with verse markup.

Use of one of following types for a **div** element is mandatory when a provided type is applicable.

- acknowledgement
- afterword
- annotant
- appendix
- article
- back
- bibliography
- body

37

- book
- bookGroup
- bridge
- chapter
- colophon
- commentary
- concordance
- coverPage
- dedication
- devotional
- entry
- front
- gazetteer
- glossary
- imprimatur
- index
- introduction
- majorSection
- map
- outline
- paragraph
- part
- preface
- publicationData
- section
- subSection
- summary
- tableofContents
- titlePage

If a type is needed but not provided, it may be added but must begin with "x-", to distinguish it from OSIS-standard values.

For example, if a user wanted to mark a **div** element with the type **reading**, the proper encoding of the start-tag for that **div** would be as follows:

```
<div type="x-reading">
```

Such markup forms the primary backbone of an OSIS document. **Chapter** and **verse** elements are important (particularly for retrieval), but considered to be an overlay onto the more linguistic or thematic structure. Given the prevalence of book/chapter/paragraph divisions in modern translations, that is the hierarchy that prevails in all cases of conflict. Most cases of conflict will be with the marking of verses. Some translations do not follow the modern practice of using paragraphs in translation, so there is no conflict between verses and paragraphs. As long as verses and chapters do not cross the boundaries of other elements, they may be expressed in as shown in the following example (NASB):

```
<chapter osisID="Mark.10">
    <title>Mark Chapter 10</title>
        <div type="section"><head>Divorce</head>
```

```
<verse osisID="Mark.10.1">Jesus then left that place and went into
the region of Judea and across the Jordan. Again crowds of people
came to him, and as was his custom, he taught them.
</verse>
<verse osisID="Mark.10.2">Some Pharisees came and tested him by
asking, "Is it lawful for a man to divorce his wife?"
</verse>
<verse osisID="Mark.10.3">"What did Moses command you?" he replied.
</verse>
<verse osisID="Mark.10.4">They said, "Moses permitted a man to write
a certificate of divorce and send her away."
</verse>
<verse osisID="Mark.10.5">"It was because your hearts were hard that
Moses wrote you this law," Jesus replied. </verse>
<verse osisID="Mark.10.6">"But at the beginning of creation God 'made
them male and female.' </verse>
<verse osisID="Mark.10.7">'For this reason a man will leave his
father and mother and be united to his wife,</verse>
<verse osisID="Mark.10.8">and the two will become one flesh.' So they
are no longer two, but one. </verse>
<verse osisID="Mark.10.9">Therefore what God has joined together, let
man not separate."</verse>
<verse osisID="Mark.10.10">When they were in the house again, the
disciples asked Jesus about this. </verse>
<verse osisID="Mark.10.11">He answered, "Anyone who divorces his wife
and marries another woman commits adultery against her. </verse>
<verse osisID="Mark.10.12">And if she divorces her husband and
marries another man, she commits adultery." </verse>
</div>
...
</chapter>
```

The quotes that appear in this example present a special problem as they do cross the boundaries of verses in an edition. Rather than introduce the mechanism for such cases here, the traditional quotation marks are used in this example. See the section *Elements that cross other elements* for details on dealing with such quotations.

In cases where the translation follows the modern practice of using paragraphs in the translation that cross verse boundaries, it is necessary that all verses be marked using the techniques described in *Elements that cross other elements*. That may seem like a burden but in order to have easy processing for a text, it is necessary that all similar parts of it be marked in the same way. As work on OSIS and other XML editors continue, it will become easier concentrate on the substance of the text and allow automatic mechanisms to deal with the technical niceties of the underlying markup. Until then, however, you will need to compensate for the weaknesses of XML processors so that your Bible text can be easily produced for the WWW, print, cellphones and other devices.

7.5.1. Scope

The **div** element has a **scope** attribute that should be used when it contains less than a complete work. For example, some Bibles only contain Proverbs and Psalms, in which case a **div** element should say, using the **scope element**, that it only contains Proverbs and Psalms. That is done using the **scope** attribute. The value of this attribute is written using the rules for osisRefs.

7.5.2. Elements allowed in div elements

The **div** element allows the following elements to occur within it:

• **a** • **abbr** • **chapter** • **closer** • **date** • **div** • **divineName** • **figure** • **foreign** • **hi** • **index** • **inscription** • **lb** • **lg** • **list** • **mentioned** • **milestone** • **name** • **note** • **p** • **q** • **reference** • **salute** • **seg** • **signed** • **speaker** • **speech** • **table** • **title** • **transChange** • **verse** • **w**

It also allows for mixed (text and element) content.

7.6. Chapter

The **chapter** element is the only **syntactic sugar** element in the OSIS schema. The term **syntactic sugar** is used in markup circles to describe an element that has been provided because it is more familiar to users. The elements that can occur in the **chapter** element are the same as those in the **div** element. It was added so that users encoding books could use it instead of **<div type="chapter">**.

The **chapter** element has a **chapterTitle** attribute that enables users to insert a text string to be used as the chapter's title.

7.6.1. Elements allowed in chapter elements

The **chapter** element allows the following elements to occur within it:

• **a** • **abbr** • **closer** • **date** • **div** • **divineName** • **figure** • **foreign** • **hi** • **index** • **inscription** • **lb** • **lg** • **list** • **mentioned** • **milestone** • **name** • **note** • **p** • **q** • **reference** • **salute** • **seg** • **signed** • **speaker** • **speech** • **table** • **title** • **transChange** • **verse** • **w**

It is a milestoneable element.

Note that it occurs only inside of a **div** element where **type="book"** and may **not** contain other **chapter** elements.

It also allows for mixed (text and element) content.

8. Paragraphs, Quotes, and Notes

Paragraphs (element **p**), quotations (element **q**), and other grouping elements can be inserted around groups of verses, as shown below. Likewise, **note** elements can be inserted where needed. The paragraph need not have an **osisID** attribute value for the set of verses it contains, since they are typically provided on the **verse** elements themselves.

8.1. p

The **p** element is the basic holder of prose in a text. As noted above, the paragraph is used in modern translations in preference to traditional verse division of the text of Bibles. There are cases where paragraphs and verses do not conflict, but there are enough cases where they do that verses should always be represented, for entire Bibles at any rate, using the milestonable version of the **verse** element. Where verses do not cross the boundaries of paragraphs, they can be represented as containers, as shown below:

```
...
<p>
<verse osisID="Esth.4.10">Then Esther spoke to Hathach, and gave him
a command for Mordecai: </verse>

<verse osisID="Esth.4.11"><q>All the king's servants and the people
of the king's provinces know that any man or woman who goes into the
inner court to the king, who has not been called, he has but one law:
put all to death, except the one to whom the king holds out the
golden scepter, that he may live. Yet I myself have not been called
to go in to the king these thirty days.</q> </verse>

<verse osisID="Esth.4.12">So they told Mordecai Esther's words.
</verse> </p>

<p>
<verse osisID="Esth.4.13">And Mordecai told them to answer Esther:
<q sID="q.Esth.4.13">Do not think in your heart that you will escape in the king's
palaceany more than all the other Jews. </verse>
</p>

<p>
<verse osisID="Esth.4.14">For if you remain completely silent at this
time, relief and deliverance will arise for the Jews from another
place, but you and your father's house will perish. Yet who knows
whether you have come to the kingdom for such a time as this?<q eID="q.Esth.4.13">
</verse>
</p>
```

Note that in this example all the paragraphs and quotations still enclose an exact number of verses. This is the **exception** rather than the rule. Caution is urged in deciding to follow this example other than for small portions of text that are known to **nest** properly. (**Nest** simply means that each element fits within the one above it, in this case, the **verse** elements "nest" within the **p** elements.

The following are some illustrations of material in a text that you would encode using the <p> element:

The Preaching of John the Baptist
(Matthew 3.1-12; Luke 3.1-18; John 1.19-28)

The first gray section would be encoded as follows:

```
...

<p>
<verse sID="Mark.1.1" osisID="Mark.1.1"/>This is the good news about
        Jesus Christ, the Son of God.<verse eID="Mark.1.1"/> <verse
        sID="Mark.1.2" osisID="Mark.1.2"/>It began just as the prophet
        Isaih has written:....<verse eID="Mark.1.2"/>
</p>
```

Note the marking in the milestoneable form of the verse element. The paragraph actually continues to the end of verse 3 but we can anticipate that other verses are going to cross boundaries and so to be consistent and conformant with OSIS, all verses should be marked using the milestoneable form of the element.

8.1.1. Elements allowed in p elements

The **p** element allows the following elements to occur within it:

• **a** • **abbr** • **catchWord** • **chapter** (only in milestone form) • **closer** • **date** • **divineName** • **figure** • **foreign** • **hi** • **index** • **inscription** • **lb** • **lg** • **list** • **mentioned** • **milestone** • **name** • **note** • **q** • **rdg** • **rdgGrp** • **reference** • **salute** • **seg** • **signed** • **speaker** • **speech** • **table** • **title** • **transChange** • **verse** • **w**

It is not a milestoneable element. It allows for mixed (text and element) content.

8.2. q

Quotations are marked using the **q** element. OSIS offers two ways to handle what quotations marks will appear in an OSIS text when **q** elements are processed. Although two separate methods are supported, texts should choose one or the other and use it consistently. One of the things that makes processing texts in markup difficult is having equally valid methods used inconsistently in the same text.

One method of handling the quotation marks that will appear in a rendered text (whether to the screen or on

paper) is to rely upon an XSLT stylesheet to insert the correct quotations marks. The rendering for quotations marks after an interruption, for example, can be distinguished using the type attribute on this element, with values such as initial, medial, and final. The stylesheet must account for embedded quotations when calculating the correct quotations marks to insert in the presentation.

Another method of handling the quotation marks to be inserted is to use the **marker** attribute on the **q** element. The encoder will insert into the value for that attribute the quotation mark that should appear in the text when it is rendered.

As noted above, it is a bad idea to mix these two methods in a single document. It is also a bad idea to use both within a quote that contains another quote. That is the encoder should not use stylesheet rendering for the outer quote (or the marker attribute) and then switch to the other method for the embedded quote. That is the sort of encoding that will make your OSIS document more difficult for others to process.

The **who** attribute of the **q** element is used where the text does not indicate the speaker, but it is important to record who is speaking. Where the speaker is named in the text, use the **<speaker>** element.

Some translations render quotes differently depending upon the speaker. A good example of that are 'red letter' editions that render the words of Jesus in red. These are all quotations and should be marked with the **q** element, so how does it get rendered as red rather than with quotation marks?

The **q** element has an attribute called **who**. In cases where it is important to mark who is being quoted in a **q** element, the encoder puts the speaker's name in the **who** attribute. For example:

```
<verse osisID="Luke.22.35 sID="Luke.22.35"/>
Then Jesus asked them, <q who="Jesus">When I sent you without purse,
    bag or sandals, did you lack anything?</q>
<verse eID="Luke.22.35"/>
```

This is a good illustration of how markup enables different renderings of a text. It is certainly possible to simply render this quote like all others in the text with traditional quotation marks. But, following the traditional rendering is often desired by publishers as well as readers.

Rendering the words of Jesus in red, such as the ones quoted above, is a matter of selecting the words of Jesus and not quotes from others for this rendering. That is the job of an XSLT stylesheet, that is the mechanism that takes your OSIS XML file and renders it to HTML, PDF or other format.

It is not uncommon for quotes to be nested inside other quotes. As an aid for automated processing of such texts, OSIS provides a **level** attribute on the **q** element to enable encoders to indicate the level of a quotation for display purposes.

An example of two quotes.

8.2.1. Elements allowed in q elements

The **q** element allows the following elements to occur within it:

• **a** • **abbr** • **closer** • **date** • **divineName** • **foreign** • **hi** • **index** • **inscription** • **lb** • **lg** • **list** • **mentioned** • **milestone** • **name** • **note** • **p** • **q** • **reference** • **salute** • **seg** • **signed** • **speaker** • **transChange** • **verse** • **w**

It is a milestoneable element. It also allows for mixed (text and element) content.

8.3. note

Many editions of the Bible have notes of several types. A number of predefined types, and some additional internal structure, are discussed later. It is customary to include the notes directly within the text, at the point to which they apply. This can be done via the **note** element, which can be placed almost anywhere.

Every **note** should have a **type** attribute to indicate its purpose; many Bible editions show different kinds of notes in different places. The pre-defined note types are listed below; they are not sharply-defined, wholly distinct categories. In addition, if none of these categories suffice, encoders may create their own so long as their names begin with "x-".

- **allusion** The note explains an implicit reference the text makes to another text or concept.
- **alternative** The note records an alternate possible reading of the text, whether due to ambiguity in translation or to manuscript variation. This type of note will almost always include an **<rdg>** element.
- **background** The note provides background information, such as cultural norms, explanations of geographic or other information original readers would have known, and so on.
- **citation** The note cites a supporting text or further explanation of some kind.
- **crossReference** The note provides a cross-reference to a related passage or other text.
- **devotional** The note includes information of interest for devotional reading.
- **exegesis** The note discusses a relevant point of exegesis or interpretation

- **explanation** The note explains implicit, ambiguous, or otherwise non-obvious aspects of the passage.
- **speaker** This type is intended mainly for use in sermons and other performance texts, where the performer may wish to make notes to him or herself. For example, "tell joke here".
- **study** The note provides helps for a deeper study of the passage.
- **translation** The note discusses an issue of translation, such as a word whose meaning is unclear in the original, or a reasons for the translator's choice of phrasing. Bible translation projects will likely use this heavily, using the subtype attribute to mark the status of each note as resolved or unresolved, the person responsible for the note, and so on.
- **variant** The note records a textual variation among source manuscripts. This will also often include an **<rdg>** element.

It is worth mentioning that although a **note** element is physically contained in the OSIS document, at the point where it applies, it is not *logically* 'contained' there. For example, a user who searches for the phrase 'undivided devotion' should find it in 1 Corinthians 7.35 of the NIV, even though there is a (cross-reference) note at the end of 'undivided,' between the words. Software that handles such cases correctly is not commonplace at this time; however, by using explicit XML markup to identify relevant components such as note, such software is made feasible.

The **note** element can appear in any element that can contain text outside of the **header** element. When you first start to use the **note** it may seem strange that the note is located 'inline' where you normally see a footnote number or other reference to help you find the note. That is a normal reaction and what is happening when you encode a text, is that you are behind the screen so to speak, before the text has been rendered. In some cases, you may want the material held in the **note** elements to be printed at the bottom of a column or page. In other cases you want to gather them up to the end of a chapter (common in some commentaries); or, for a Sunday School class, you may want to suppress them altogether. Once your text is in OSIS markup, all of those possibilities and more are open to you.

The following is a fairly complex example of different types of notes in a text. The encoded text is from the *Contemporary English Version (CEV)*, the Book of Esther, chapter 1, verses 1 and 2. The example is explained line by line following the example.

1. <verse sID="Esth.1.1-Esth.1.2" osisID="Esth.1.1 Esth.1.2" n="1-2"/>
2. <note type="crossReference" osisRef="Esth.1.1-Esth.1.2" osisID="Esth.1.1!crossReference">
3. <reference osisRef="Ezra.4.6">Ezra 4:6</reference></note>
4. King Xerxes
5. <note type="translation" osisRef="Esth.1.1-Esth.1.2" osisID="Esth.1.1!note.a" n="a">
6. <catchWord>Xerxes</catchWord>The Hebrew text has <q level="1">Ahasuerus,</q> who was better known as King Xerxes I (485-465 B.C.).</note>
7. of Persia lived in his capital city of Susa
8. <note type="alternate" osisRef="Esth.1.1-Esth.1.2" osisID="Esth.1.1!note.b" n="b">
9. <catchWord>in his capital city of Susa</catchWord> Or <q level="1">in his royal fortress in the city of Susa.</q> Susa was a city east of Babylon and a winter home for Persian kings.</note>
10. and ruled one hundred twenty-seven provinces from India to Ethiopia.
11. <note type="background" osisRef="Esth.1.1-Esth.1.2" osisID="Esth.1.1!note.c" n="c"> <catchWord>Ethiopia</catchWord>The Hebrew text has <q level="1">Cush,</q> which was a region south of Egypt that included parts of the present countries of Ethiopia and Sudan.</note>
12. <verse eID="Esth.1.1-Esth.1.2"/>

Now for the explanation:

1. First line: <verse sID="Esth.1.1-Esth.1.2" osisID="Esth.1.1 Esth.1.2" n="1-2"/>

   This line starts the first verse of the first chapter of the Book of Esther. Note that it uses the milestoneable version of the **verse** element. A **milestoneable** element is one that can be written with a start-tag, **<verse>** and an end-tag, **</verse>**, but can also be written with two milestones, thus, **<verse sID="1"/>** and **<verse eID="1"/>**, where the **sID** and **eID** values link the two milestone elements together. See *Elements that cross other elements* for more details on this OSIS mechanism. Here, the verse element includes what was in the KJV, for example, written as Esther 1:1 and Esther 1:2, it has the **sID** attribute with a range value of "Esth.1.1-Esth.1.2" to match the **eID** attribute you see on line 11.

   It also has an **osisID** attribute that indicates this verse contains both "Esth.1.1" and "Esth.1.2." There is only one **osisID** attribute but the verses identified by it are both listed, separately. You cannot have a range, like, Esth.1.1-2 for an **osisID**. Ranges are permitted for **osisRef** attributes, which we will meet on the next line.

   The final attribute you see is the **n** attribute. This is most commonly used to record some numbering that was present in an original text that may or may not have any relationship to the structure being encoded. In other words, the CEV translations inserted "1-2" here and OSIS provides a mechanism to record the fact of "1-2" occuring here.

2. Second line: <note type="crossReference" osisRef="Esth.1.1-Esth.1.2" osisID="Esth.1.1!crossReference">

   The second line of the encoding has not (yet) reached the text of Esther. The translators have inserted a note to direct the reader to other material. This is marked using the **note** element with several attributes. First, there is the **type** attribute that records this note is a cross reference to another text, using the **type** of **crossReference**. (Recall our earlier discussion about XML being case sensitive? Well, "cross reference," "crossreference" and "CrossReference" will all cause errors if used. Use "crossReference" only. It is written in "camel case," that is the first letter of the first word is in lower case and the beginning of each following word (no spaces) is in upper case. Another tidbit for the church picnic.)

   The next thing to notice is that the **note** has an **osisRef** attribute and value. The value **Esth.1.1-Esth.1.2** means that this **note** is pointing at the material in this text that has the **osisID** of **Esth.1.1-Esth.1.2**. In practice this means that no matter where these notes are placed, physically speaking, they are always "pointing" at this particular place in the text. It is, however, a **best practice** to put notes where they occur in the text.

   The third thing to notice is that the note has an **osisID** attribute and value. **Note** elements are not required to have an **osisID** but in order to link to a particular note from other notes or other parts of the text, an **osisID** is necessary. It should be the exceptional case when a note does not have an **osisID**.

   Finally, notice that this **note** does not have an **n** attribute. The other notes in this example do. The reason is that cross references are often just printed as just the reference, without any note number. Here the text for that would be drawn from the **osisRef** attribute value.

3. Third line: <reference osisRef="Ezra.4.6">Ezra 4:6</reference></note>

This line starts inside the **note** element itself; this is the content of the note as it appeared in the CEV. In a printed text, one would see Ezra 4:6, but the OSIS markup provides greater abilities from the same information. The **reference** element, that is inside the note, has an **osisRef** attribute. That attribute points at Ezra 4:6, but written as Ezra.4.6. The reason for that pointing is to allow applications to present the reader with a hypertext link to that text or even display it automatically (if the user so chooses) along with the main text. Note that the traditional rendering of the Ezra citation is also preserved for the reader.

4. Forth line: King Xerxes

   Finally, the first word in the text! It does seem like a lot to get this far, but recall that the encoding compensates for how dumb computers are as compared to a human reader. The advantages for publishing, distribution and study are well known, so lets see if we can get to the end of the first two verses!

5. Fifth line: <note type="translation" osisRef="Esth.1.1-Esth.1.2" osisID="Esth.1.1!note.a" n="a">

   Arrgh! Another note! Well, any well translated Bible is going to have lots of notes so we start with an example that has four of them.

   The second **note** element is different from the first one. Can you see how? This **note** element has a different **type** attribute, that of **translation**. Translation notes usually indicate additional information for the reader to consider in evaluating the translation.

   Remember that the **canonical** attribute for the **note** element defaults to false. That is to say that the content of the note is not part of the biblical text.

   Note that the range of an **osisRef** for a note can overlap other notes. In other words, there could be separate notes that had Gen.1.5 and Gen.1.5-7 as their **osisRef** values.

   Finally, see that this note has the **n** attribute with a value of "a." That is the text that would appear in the text preceding note. That value could also be derived from the **osisID** for the note. Which on is used marks no difference but again, encoders should be consistent in a single document.

6. Sixth line: <catchWord>Xerxes</catchWord>The Hebrew text has <q level="1">Ahasuerus,</q> who was better known as King Xerxes I (485-465 B.C.).</note>

   Inside this **note** element, we encounter the **catchWord** element. The **catchWord** element contains text that occurs in the main text and works like a pointer for a human reader as to which word or phrase the **note** applies to in the text. In this case, the note is concerned with the name, "Xerxes," which we just encountered as the first word in the text. The note goes on to inform the reader that the Hebrew text differs from that which appears in the translation, along with an explanation for that difference.

   This **note** element also contains a **q** element that has quoted material that appears in the note. The **q** element has a **level** attribute that is used to indicate the type of quotation marks that should appear in rendering this particular quote. Think of it as being an indication of where you have a quote that is inside another quote. Normally the outside quote has double quotation marks, while the inner quote has only single quotation marks. The **level** attribute is used to preserve that sort of rendering information, when automatic rendering is not used.

7. Seventh line: of Persia lived in his capital city of Susa

   Well! Another line of text!

8. Eight line: <note type="alternate" osisRef="Esth.1.1-Esth.1.2" osisID="Esth.1.1!note.b" n="b">

   The third **note** element to appear for this passage. It has all the same attributes as those covered at line 5, so that discussion is not repeated here.

9. Ninth line: <catchWord>in his capital city of Susa</catchWord> Or <q level="1">in his royal fortress in the city of Susa.</q> Susa was a city east of Babylon and a winter home for Persian kings.</note>

   The content of the third **note** is similar to that covered at line 6. Note that the content of **catchWord** can be an entire phrase.

10. Tenth line: and ruled one hundred twenty-seven provinces from India to Ethiopia.

    The last line of text that we will encounter here.

11. Eleventh line: <note type="background" osisRef="Esth.1.1-Esth.1.2" osisID="Esth.1.1!note.c" n="c"> <catchWord>Ethiopia</catchWord>The Hebrew text has <q level="1">Cush,</q> which was a region south of Egypt that included parts of the present countries of Ethiopia and Sudan.</note>

    The fourth (and final) **note** element in this passage. See lines 5 and 6 for the meaning of the various parts.

12. Twelfth line: <verse eID="Esth.1.1-Esth.1.2"/>

    The milestoneable **verse** element that marks the end of this passage. Note that the **eID** value matches that found at line 1.

That may seem like a lot of markup but look at the explanation again. Note how as we moved through the markup, the less and less explanation was needed? That is because once you master some basic elements and concepts of markup, you can combine those to make a richly encoded text that will serve many purposes.

8.4. Note Placement

The **placement** attribute on the **note** element is used to specify the placement of a note within the text. In addition to the usual "x-" preceeding a custom value, the following values have been defined:

- foot
- end
- inline
- left
- right
- interlinear
- apparatus

8.5. Examples of Notes

8.5.1. Alternative Readings

ᶜ we need; or for today, or for tomorrow.

An alternative reading in a note.

ıme   and went with him.

of God.   1:2 Mal 3:1.   1:3 Isa 40:3.   1:4 Greek *preaching a*
1:7 Greek *to stoop down and untie his sandals.*   1:8 Or *in.*

Another alternative reading in a note.

8.5.2. Notes with <catchWord>

Moses . . . well   Read about this in Leviticus 14:1-32.

The bold text is encoded with the <catchWord> element to show that it appears in the main text. Only the first and last word of the phrase are shown.

ᶜ **we need;** or for today, or for tomorrow.

The bold text here is also encoded with a <catchWord> element. The following **for today**, **for tomorrow** should be encoded with the <rdg> element.

8.5.3. Note Markers

The Preaching of John the Baptist
(Matthew 3.1-12; Luke 3.1-18; John 1.19-28)

1 This is the Good News about Jesus Christ, the Son of God.ᵃ ²It began as the prophet Isaiah had written:
"God said, 'I will send my
        messenger ahead of you
    to open the way for you.'
³Someone is shouting in the desert,
    'Get the road ready for the
        Lord;
    make a straight path for him to
        travel!' "
⁴So John appeared in the desert, baptizing and preaching.ᵇ "Turn away from

The grayed area is a marker that appears in the text. Note this is not encoded as part of the note.

> 66 Iri bukatejuukana*, abashaaja ba'Bayahudi banagira inaama kuguma na'bakulu ba'bagingi, na'bigiriza be'maaja. Ne'ri Yesu akaleetwa i'mbere lyabo, 67 banamúbuuza: «Iri wangaba we'Kiri-sito, utubwire.» Yesu anashuvya: «Kundu nangamùbwira, mutangayemeera. 68 Na'kundu nangamùbuuza igambo, mutanga-njuvya. 69 Haliko ukulyokera buno, Umwana wo'Mundu agaba abwatiiri i'luhande lwe'lulyo k lwa'Rurema wo'bushobozi bwoshi.»
>
> 70 Yabo booshi banashubi buuza Yesu: «Ka'kwokwo uli Mwana wa'Rurema?» Anabashuvya: «Si'mwenyene keera mwadeta kwo'ndi yehe!» 71 Lyeryo yabo bakulu banadeta: «Aaho!
>
> k22:69 Ukubiikwa uluhande lwe'lulyo, kuli kuhaabwa ubushobozi no'lushaagwa.

Another example of a note marker. Note markers are produced by the stylesheet that is used with an OSIS document. This allows large print markers for visually impaired users or other forms of presentation.

Note markers can also be preserved by using the **n** attribute on the **note** element but that is not recommended.

8.5.4. Notes with References

> **The Ancestors of Jesus Christ**
> *(Luke 3.23–38)*
>
> 1 This is the list of the ancestors of Jesus Christ, a descendant of David, who was a descendant of Abraham.
>
> 2–6a From Abraham to King David, the following ancestors are listed: Abraham, Isaac, Jacob, Judah and his brothers; then Perez and Zerah (their mother was Tamar), Hezron, Ram, Amminadab, Nahshon, Salmon, Boaz (his mother was Rahab), Obed (his mother was Ruth), Jesse, and King David.
>
> 1.11 2 K 24.14–15; 2 Ch 36.10; Jr 27.20.

This note contains references that are encoded using the <reference> element.

> ıme   and went with him.
>
> of God.   1:2 Mal 3:1.   1:3 Isa 40:3.   1:4 Greek *preaching a*
> 1:7 Greek *to stoop down and untie his sandals.*   1:8 Or *in.*

Another example where <reference> elements should be used.

50

8.5.5. Notes with Verses

> **Verse 6**   Some Greek copies add 6b-8a: "And we wanted to judge him by our own
> used much force to take him from us. ⁸And Lysias commanded those who wanted to

A fairly complex note, which has a <p> element, which contains a <verse> element. Note that the content of the <note> element itself does not use the <p> element.

8.6. Elements allowed in note elements

The **note** element allows the following elements to occur within it:

• **a** • **abbr** • **catchWord** • **date** • **divineName** • **figure** • **foreign** • **hi** • **index** • **inscription** • **lb** • **lg** • **list** • **mentioned** • **milestone** • **name** • **note** • **p** • **q** • **rdg** • **rdgGrp** • **reference** • **seg** • **table** • **title** • **verse** • **w**

It is not a milestoneable element. It does allows for mixed (text and element) content.

9. Interim Summary

At this point you have seen four (4) elements that you will using everyday. Namely, **div**, **p**, **q**, and **note**. Assuming someone prepares the header for the document for you, you are well on your way to doing serious encoding of a Bible.

Granted, you are going to want to mark names, poetry, lists, insert tables and illustrations and OSIS supports all of that and more. But, the important point is that by knowing the elements that you need to use everyday, which is far less than even the entire set provided by OSIS, you can do professional quality work on encoding biblical texts.

There is one issue that needs to be covered before going further on OSIS markup. With the exception of the **verse** elements in the last example, all of the elements were traditional containers. That is to say, an element opens, contains text and perhaps other elements, then closes, fully surrounding all the text and elements inside. But, that is not always the case, particularly in Bible translations.

Sometimes a **verse** or **chapter** starts or end in the middle of some other unit, such as a poetic line group, paragraph, quotation, or speech. In such cases an alternate form of the **verse** or **chapter** tags must be used.

OSIS provides a very robust mechanism for handling such cases and it is based upon the elements you are **already learning**. That is to say that you won't have to learn new elements in addition to the ones you use everyday for encoding Bibles. You will need to learn only two more attributes for each element and they will be easy to remember. One is for the **start** of an element and the other is for the **end** of the same element.

10. Elements that cross other elements

The normal form of an element is a start tag and an end tag: **\<q>...\</q>**. For handling markup that crosses boundaries, however, a special form must be used. It consists of two totally empty instances of the same element type: one to mark the starting point, and one to mark the ending point. The two empty elements identify themselves as to which is the start and which is the end, and co-identify themselves by the **sID** attribute (the start of the traditional element) and the **eID** attribute (the end of the traditional element), the values of which must match.

Empty elements are indicated in XML by a tag with "/" preceding the final ">": thus **\<q/>** rather than **\<q>** or **\</q>**. Elements used in this way are commonly called 'milestones,' and those particular elements in OSIS that permit this alternate encoding are thus called 'milestoneable.' Elements that are 'milestoneable' in the OSIS schema are:

- **abbr**
- **chapter**
- **closer**
- **foreign**
- **l**
- **lg**
- **q**
- **salute**
- **seg**
- **signed**
- **speech**
- **verse**

This is particularly useful where modern translations break up verses or other traditional divisions in a Bible text. For example, a paragraph based encoding of part of the Book of Esther would appear as follows:

```
<p>
<verse sID="Esth.2.7" osisID="Esth.2.7"/>Mordecai had a very beautiful cousin named
Esther, whose Hebrew name was Hadassah. He had raised her as his own daughter, after
her father and mother died.
<verse eID="Esth.2.7"/>

<verse sID="Esth.2.8" osisID="Esth.2.8"/>When the king ordered the search for
beautiful women, many were taken to the king's palace in Susa, and Esther was one of
them.
</p>
<p>Hegai was put in charge of all the women,
<verse eID="Esth.2.8"/>

<verse sID="Esth.2.9" osisID="Esth.2.9"/>and from the first day, Esther was his
favorite. He began her beauty treatments at once. He also gave her plenty of food
and seven special maids from the king's palace, and they had the best rooms.
<verse eID="Esth.2.9"/>
</p>
```

There are two things to note about the Esther example:

- Esther 2:8 is divided by a paragraph (the **p** element) and so must be marked using the **verse** element as a milestones with the **sID** and **eID** attributes to link those two milestones together.
- Where overlapping elements are necessary, the milestoneable element technique must be used for the entire text. That is, it is an error to mark some verses in Esther with traditional **verse** elements, i.e., as containers and others with the milestoneable verses. The reason is quite simple, inconsistent markup is more difficult to process and makes the encoded text less useful for everyone.

This is equivalent to the TEI "milestone" method for marking such phenomena. It has the advantage that milestones representing a given type of element have the same name as the element, and automatically have the same attributes. Standard XML processors will not detect a validation error if attributes other than **eID** are specified on the ending **milestone**, if the **eID** is specified on the starting **milestone**, or the start and end milestones are in the wrong order, each of these conditions is an OSIS error.

For OSIS purposes, there is no semantic difference between marking up a chapter or verse as a container using a start and end tag, versus marking it up as a "milestone pair" consisting of two empty tags.

Note: Typesetting and layout systems vary in their ability to accommodate non-hierarchical markup such as this. Fortunately, in most Bible editions the only formatting consequence of a **verse** element is insertion of the verse number, and perhaps insertion of a line-break; these are within the capabilities of most layout and style systems even though the verse is not a container in XML terms.

While all the elements listed above can be written as milestones, the most common ones will be **l**, **q**, **speech**, and **verse**.

For future reference: An element in its milestoneable form should end without crossing boundaries not required to mark the element. For example, if the milestone form of **verse** is being used, it is not an XML error to put the ending milestone outside of a closing **div** element, but if not required by the content of the **verse** element, it is an OSIS error.

To illustrate that point, consider the following examples. The first one is **incorrect** and the second one is **correct**.

Incorrect Usage

```
<p>
<verse sID="Esth.2.7" osisID="Esth.2.7"/>Mordecai had a very beautiful cousin named
Esther, whose Hebrew name was Hadassah. He had raised her as his own daughter, after
her father and mother died.
</p>
<verse eID="Esth.2.7"/>
```

Note that the closing element does not enclose any more text but does enclose the closing tag for the **p** element. There is no advantage to this approach and it causes difficulty for most XML processing solutions.

Correct Usage

```
<p>
<verse sID="Esth.2.7" osisID="Esth.2.7"/>Mordecai had a very beautiful cousin named
Esther, whose Hebrew name was Hadassah. He had raised her as his own daughter, after
her father and mother died.
<verse eID="Esth.2.7"/>
</p>
```

Note that the closing element encloses the same text as before, but now it is inside the closing tag for the **p** element. This simplifies processing of OSIS documents and is likely to be specified formally in a future OSIS release.

11. Special Text Types

The bulk of the remaining OSIS elements fall into a few simple classes: First, markup for special text types, such as epistles and drama. Second, generic structures such as lists, tables and glossaries (typically found in appendixes of printed Bibles). And finally, small-scale elements that mark quotations, notes, names, index entries, and the like.

11.1. Markup for epistles and similar materials

Letters, epistles, and similar texts are marked up in basically the same way as any other text. However, three special elements are available for marking portions unique to this genre of text:

11.1.1. salute

The **salute** element encloses the salutation or greeting, typically at the very beginning of a letter. It should include the whole salutation, including (if present) the "to", "from", and any following greeting or blessing. If the boundaries of a salutation are the same as the boundaries of a paragraph, section, or other unit, that unit should be placed outside, with the **salute** element directly within. For example (LBP):

```
<div type="book" osisID="1Tim">
    <title>The First Epistle to Timothy</title>
    <chapter osisID="1Tim.1">
       <salute>
          <verse osisID="1Tim.1.1" sID="1Tim.1.1"/>FROM: PAUL, a missionary of Jesus
          Christ, sent out by the direct command of God our Savior and by Jesus
          Christ our Lord -- our only hope.<verse eID="1Tim.1.1"/>
          <verse osisID="1Tim.1.2" sID="1Tim.1.2"/>To: Timothy. Timothy, you are
          like a son to me in the things of the Lord. May God our Father and Jesus
          Christ our Lord show you his kindness and mercy and give you great peace
          of hear and mind.<verse eID="1Tim.1.2"/>
       </salute>
       <verse osisID="1Tim.1.3" sID="1Tim.1.3"/>...<verse eID="1Tim.1.3"/>
        </chapter>
     ...
  </div>
```

11.1.1.1. Elements allowed in salute elements

The **salute** element allows the following elements to occur within it:

• **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **lg** • **list** • **mentioned** • **milestone** • **name** • **note** • **p** • **q** • **reference** • **salute** • **seg** • **speaker** • **transChange** • **verse** • **w**

It is a milestoneable element. It also allows for mixed (text and element) content.

11.1.2. signed

The **signed** element surrounds the name of the author and/or amanuensis of a letter and its immediately surrounding phrase of opening or closing (if any). In Biblical epistles, it is common for the author to be

named only at the beginning; this should still be marked up with the **signed** element.

The **signed** element may appear with or without **closer** or **salute** elements, and the name may or may not also be tagged as a **name** (if it is, the **name** should be the inner element even if it includes all the text content of the **signed** element). In New Testament epistles, there is not generally an obvious, final signature. However, this element may be used somewhat more broadly of a phrase or portion judged as intended to identify the writer. As shown below, the signature of a secretary (amanuensis) may also be marked up in this way. For example:

- <verse osisID="Rom.16.22"><signed>I Tertius salute you which wrote this epistle in the Lorde.</signed></verse>

  [English, Tyndale, 1525/1530]

- <verse osisID="1Cor.16.21"><signed>I, Paul, write this greeting with my own hand.</signed></verse>

  [English, RSV]

- <verse osisID="2Cor.1.1"><signed>Paul, an apostle of Jesus Christ by the will of God, and Timothy [our] brother, to the church of God which is at Corinth, with all the saints who are in all Achaia:</signed></verse>

  [English, Webster]

- <verse osisID="Gal.6.11"><signed>See with what large letters I am writing to you with my own hand.</signed></verse>

  [English, RSV]

- <verse osisID="Eph.1.1"><signed>Paul, an apostle of Christ Jesus through the will of God, to the saints that are at Ephesus, and the faithful in Christ Jesus:</signed></verse>

  [English, American Standard Version, 1901]

- <verse osisID="1Thess.1.1"><signed>Paul, and Silvanus, and Timothy, to the church of the Thessalonians which is in God the Father and in the Lord Jesus Christ: Grace to you, and peace. </signed></verse>

  [English, RKJNT]

- <verse osisID="1Tim.1.1"><signed>Paul, an apostle of Jesus Christ, according to the commandment of God our Savior, and of Christ Jesus our hope:</signed></verse>

  [English, Douay-Rheims Bible, Challoner Revision]

- <verse osisID="Phm.1.1"><signed>Mimi Paulo, mfungwa kwa ajili ya Kristo Yesu, na ndugu Timotheo,</signed> ninakuandikia wewe Filemoni mpendwa, mfanyakazi mwenzetu</verse> <verse osisID="Phm.1.2">na kanisa linalokutana nyumbani kwako, na wewe dada Afia, na askari mwenzetu Arkupo.</verse>

[Swahili NT]

11.1.2.1. Elements allowed in signed elements

The **signed** element allows the following elements to occur within it:

• **a** • **abbr** • **divineName** • **foreign** • **hi** • **index** • **lb** • **milestone** • **name** • **note** • **reference** • **seg** • **w**

It is a milestoneable element. It also allows for mixed (text and element) content.

11.1.3. closer

The **closer** element surrounds the closing portion of a letter, typically consisting of final greetings or blessing and a signature (see **signed**). It is a matter of judgement just where a **closer** begins and ends. For example:

- <verse osisID="1John.5.21"><closer>Dear children, keep away from anything that might take God's place in your hearts. Amen. Sincerely, <signed>John</signed></closer></verse>

  [LBP]

11.1.3.1. Elements allowed in closer elements

The **closer** element allows the following elements to occur within it:

• **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **lg** • **list** • **mentioned** • **milestone** • **name** • **note** • **p** • **q** • **reference** • **seg** • **signed** • **speaker** • **transChange** • **verse** • **w**

It is a milestoneable element. It also allows for mixed (text and element) content.

11.1.4. Benedictions

OSIS presently provides no special markup for benedictions and blessings. Recommended practice at this time if an encoder wishes to identify them in a text, is to use **<seg type="benediction">**. For example:

- <verse osisID="2Cor.13.14"><seg type="benediction">The grace of the Lord Jesus Christ, and the love of God, and the communion of the Holy Spirit, <transChange type="added">be</transChange> with you all. Amen.</seg></verse>

11.2. Dramatic texts, castList, castItem, actor, role, roleDesc

OSIS provides two ways to encode dramatic texts such as plays. The first way is to declare the list of characters, or **castList**; the second is to identify speeches and their speakers in the body of a dramatic text.

A **castList** element contains a structured list of the roles, or cast, of a dramatic work. It is drawn directly from the TEI structure for the same thing. For example, in the Song of Songs, some translations may present the list of characters at the start of the book: lover, beloved, and friends. A similar approach might be taken

for Job. However, these elements will be most commonly used for extra-Biblical materials, such as a play based on the Bible, or dramas in classical or other literature.

A **castList** element for a dramatic re-enactment of Job would be encoded as follows:

```
<castList>
    <castGroup>
        <head>Cast of characters</head>
        <castItem>
            <actor>Patrick Durusau</actor>
            <role>Job</role>
            <roleDesc>A man of God who suffers greatly</roleDesc>
        </castItem>

        <castItem>
            <actor>(a whirlwind)</actor>
            <role>God</role>
            <roleDesc>The Almighty, who permits Job's suffering, and
            responds to his questions about it.</roleDesc>
        </castItem>

        <castItem>
            <actor>(a disembodied voice)</actor>
            <role>Satan</role>
            <roleDesc>The instigator of Job's suffering</roleDesc>
        </castItem>

        <castItem>
            <actor>Todd Tillinghast</actor>
            <role>Eliphaz</role>
            <roleDesc>The first of Job's friends to speak</roleDesc>
        </castItem>

        <castItem>
            <actor>Chris Little</actor>
            <role>Bildad</role>
            <roleDesc>The second of Job's friends to speak</roleDesc>
        </castItem>

        <castItem>
            <actor>Steve DeRose</actor>
            <role>Zophar</role>
            <roleDesc>The third of Job's friends to speak</roleDesc>
        </castItem>

        <castItem>
            <actor>Troy Griffiths</actor>
            <role>Elihu</role>
            <roleDesc>The youngest and last of Job's friends to speak,
            who was slightly less clueless than the rest.</roleDesc>
        </castItem>
    </castGroup>
</castList>
```

The **castList** element contains the entire casting list, and consists of one or more **castGroup** elements. Multiple **castGroup** elements, each with its own head, would be used if there were multiple sub-groups of the cast to be listed separately; more typically there will be only one **castGroup** within a **castList**.

At this time, **castList** can only occur in a **work** declaration, after the Dublin Core elements. Thus, if a Bible

encoder wishes to include the cast of Song of Songs and that of Job, they would each need to be marked as a separate **castGroup** within that one **castList**.

It is strongly recommended that each **castGroup** and **castItem** have an **ID** attribute. Since IDs must be unique across all element types in a document, encoders may wish to prefix certain kinds of IDs to separate them and avoid conflicts. For example, an appropriate ID for a **castItem** representing the Friends in Song of Songs would be "cast.friends", or perhaps "cast.song.friends".

11.2.1. Elements allowed in castList elements

The **castList** element allows the following elements to occur within it:

- **castGroup**

It is not a milestoneable element. It only has the **castGroup** element as its content.

11.2.2. castGroup

The **castGroup** element is the container for **castItem** elements. It may begin with an optional **head** element.

11.2.2.1. Elements allowed in castGroup elements

The **castGroup** element allows the following elements to occur within it:

• **head** • **castItem**

It is not a milestoneable element. It only allows **head** and **castItem** elements as its content.

11.2.3. castItem

The **castItem** element contains the full information for a single character. This must include a name for the **role** being played, and should include a **roleDesc**, that is, a description of that role. It may also include the name of an **actor**, if the text being encoded represents a particular enactment rather than, say, a libretto or script.

11.2.3.1. Elements allowed in castItem elements

The **castItem** element allows the following elements to occur within it:

• **actor** • **role** • **roleDesc**

It is not a milestoneable element. It only allows **actor**, **role**, and **roleDesc** elements as its content.

11.2.4. actor

In general there is no need to encode an actor name or role name with an explicit **name** element, unless the encoder wishes to provide a normalized form for later reference; in that case, the **name** element should be placed just *within* the **actor** or **role** element, not surrounding it.

This element has a **who** attribute that can be used to specify a uniform name for the actor.

11.2.4.1. Elements allowed in actor elements

The **actor** element allows the following elements to occur within it:

• **a** • **abbr** • **foreign** • **index** • **note** • **reference** • **seg** • **w**

It is not a milestoneable element. It also allows for mixed (text and element) content.

11.2.5. role

The **role** element is used to identify the role that a particular actor plays.

11.2.5.1. Elements allowed in role elements

The **role** element allows the following elements to occur within it:

• **a** • **abbr** • **foreign** • **index** • **note** • **reference** • **seg** • **w**

It is not a milestoneable element. It also allows for mixed (text and element) content.

11.2.6. roleDesc

The **roleDesc** element contains a description of the role played by a particular actor.

11.2.6.1. Elements allowed in roleDesc elements

The **roleDesc** element allows the following elements to occur within it:

• **a** • **abbr** • **foreign** • **index** • **note** • **reference** • **seg** • **w**

It is not a milestoneable element. It also allows for mixed (text and element) content.

11.3. speaker

The **speaker** element is used to encode the person or role that is uttering the content of an associated speech. A formal pointer can be included, connecting the **speaker** element to a **castItem** in the header. For example:

```
<div osisID="NRSV.Song.2">
<speech>
        <speaker who="castList1:woman">woman</speaker>
<verse osisID="NRSV.Song.2.1">I am a rose of Sharon, a lily of the valleys.</verse>
</speech>
</div>
```

Which is the equivalent to:

```
<div osisID="NRSV.Song.2">
<speech who="castList1:woman">woman
<verse osisID="NRSV.Song.2.1">I am a rose of Sharon, a lily of the valleys.</verse>
</speech>
</div>
```

Either method is correct, but a given text should be consistent in using one method or the other. It is not required that the who attribute be provided, or that it have a prefix tieing it to a **castList**; however, this is the recommended best practice because it enables computer software to assist with error-checking.

It should be noted that the **speaker** element should be used when the speaker is actually named in the text. Otherwise, the **who** attribute on **q** or **speech** should be used to hold the speaker information. The distinction is between cases where the speaker is actually named in a text versus being inferred by the reader.

The naming of a speaker will be more common in dramatic than in biblical texts. The important point is to use the **who** attribute on **q** and **speech** elements when the speaker is not named.

Note: The word *woman* does not appear in the Hebrew text as was added by the translator. To indicate that fact, the following encoding could be used:

```
<div osisID="NRSV.Song.2">
<speech>
        <transChange type="added">woman</transChange>
<verse osisID="NRSV.Song.2.1">I am a rose of Sharon, a lily of the valleys.</verse>
</speech>
</div>
```

The **workPrefix** element added in OSIS 2.1 may be used to specify the default prefix for the **who** attribute on **speech**, by specifying **path='//speech/@who'**.

11.3.1. Elements allowed in speaker elements

The **speaker** element allows the following elements to occur within it:

• **a** • **abbr** • **divineName** • **foreign** • **index** • **name** • **note** • **reference** • **w**

It is not a milestoneable element. It does allow for mixed (text and element) content.

11.4. speech

The **speech** element is used to indicate quoted direct speech. In that sense it represents a kind of quotation.

However, the **q** element is to be used for quotations in general, where the **speech** element is limited to accounts of an individual making an actual speech in some kind of performance context. In general, both elements should not be applied to the same text portion. Just as with the **q** element, the marks associated with a speech can be automatically rendered or recorded using the **marker** attribute. For example:

```
<chapter osisID="Acts.7">
<title>Stephen's Speech to the Sanhedrin</title>

<verse osisID="Acts.7.1" sID="a71"/>Then the high priest asked him, <speech
who="high priest">Are
these charges true?</speech>
<verse eID="a71"/>

<verse osisID="Acts.7.2" sID="a72"/>To this he replied:
<speech who="Stephan">Brothers and fathers, listen to me! The God of glory appeared
to our father Abraham while he was still in Mesopotamia, before he lived in Haran
<verse eID="a72"/>

<verse osisID="Acts.7.3" sID="a73"><q type="embedded" marker="'" who="God">Leave
your country and your people,</q> God said, <q type="embedded" marker="'"
who="God">and go to the land I will show you.</q><verse eID="a73"/>

<verse osisID="Acts.7.4" sID="a74"/><milestone type="cQuote" marker='"'">So he left
the land of the Chaldeans and settled in Haran. After the death of his father, God
sent him to this land where you are now living. <verse eID="a74"/>

<verse osisID="Acts.7.5" sID="a75"/>He gave him no inheritance here, not even a
foot of ground. But God promised him that he and his descendants
after him would possess the land, even though at that time Abraham
had no child. <verse eID="a75"/>

<verse osisID="Acts.7.6" sID="a76"/>God spoke to him in this way: <q
type="embedded" marker="'">Your descendants will be strangers in a country not
their own, and they will be enslaved and mistreated four hundred years. <verse
eID="a76"/>

<verse osisID="Acts.7.7" sID="a77"/>But I will punish the nation they serve as
slaves,</q> God said, <q type="embedded" marker="'">and afterward they will come out
of that country and worship me in this place.</q><verse eID="a77"/>

<verse osisID="Acts.7.8" sID="a78"/>Then he gave Abraham the covenant of
circumcision. And Abraham became the father of Isaac and circumcised
him eight days after his birth. Later Isaac became the father of
Jacob, and Jacob became the father of the twelve
patriarchs.<verse eID="a78"/>

...

<verse osisID="Acts.7.53" sID="a79"/>you who have received the law that was put
into effect through angels but have not obeyed it.
<verse eID="a79"/>
</speech>
...</chapter>
```

Note that in this example the high priest's short speech in verse 1 is marked up as a normal container element with normal start- and end-tags, as is Stephen's reply. But, note that all the verse boundaries have been represented with milestoneable verse elements. The reason for this is quite simple, if the encoding jumps from using containers for verses and only on occasion changes to milestones, noting that Stephen's speech start inside a verse, the file becomes very difficult to process reliably. When a conflict arises between the

scope of chapter/verse units and other units, the chapter/verse units give way by being represented as milestones. If a conflict arises between two other units (say, a quote that encompasses part but not all of each of two paragraphs), it is left to the encoder's discretion which of them is represented via milestones.

11.4.1. Elements allowed in speech elements

The **speech** element allows the following elements to occur within it:

• **a** • **abbr** • **chapter** • **closer** • **date** • **divineName** • **foreign** • **hi** • **index** • **inscription** • **lb** • **lg** • **list** • **mentioned** • **milestone** • **name** • **note** • **p** • **q** • **reference** • **salute** • **seg** • **signed** • **speaker** • **table** • **title** • **transChange** • **verse** • **w**

It is a milestoneable element. It also allows for mixed (text and element) content.

11.5. Marking up poetic material

Although poetic material is commonly called "verse" material, OSIS avoids that term because of potential confusion with the book/chapter/verse reference system. Thus, as in TEI, markup of poetry refers to lines and line groups.

OSIS provides a typographic line-break element, namely **lb**. Although treated under poetry, note that the lb element is generally available for use in other elements. This should aid in the transformation of prior Bible encoding systems into OSIS.

These three elements for marking up poetic material work as follows:

11.5.1. lg

The **lg** or "line group" element is used to contain any group of poetic lines. Thus it covers for units like couplet, stanza, and entire poem; these may be distinguished via the type attribute.

No pre-defined types are provided; any types used should begin with "x-" so as not to conflict with standard values that may be agreed upon later and added to OSIS.

The **lg** element can contain other **lg** elements that contain smaller groups of lines.

11.5.1.1. Elements allowed in lg elements

The **lg** element allows the following elements to occur within it:

• **chapter** • **index** • **l** • **lb** • **lg** • **milestone** • **q** • **verse**

It is a milestoneable element. In its container form, it allows only for element content.

11.5.2. l

The **l** element is used to mark poetic lines, as determined by the linguistic nature of poetry in the language of the work. For example, much English poetry consists of lines that can be located by the position of rhyming words, and/or by counting syllables; Hebrew poetry can often be divided into lines based on parallelism of thought or meaning.

The **level** attribute on the **line** element must start with the number 1 and increase only by whole numbers.

The following example shows an encoding of the first two verses of Psalm 7 from the CEV which uses the **lg** and **l** elements to mark poetic material.

```
<div type='section' scope='Ps.7.1-Ps.7.17'>
      <title>The <divineName type='x-yhwh'>LORD</divineName> Always Does
       Right</title>
   <lg>
     <l>
         <verse sID='Ps.7.1' osisID='Ps.7.1'/>You, <divineName
          type='yhwh'>LORD</divineName> God,
          <lb type='x-secondLine'/>are my protector.
     </l>
     <l>Rescue me and keep me safe<lb type='x-secondLine'/>from all who chase
        me.<verse eID='Ps.7.1'/>
     </l>
     <l>
         <verse sID='Ps.7.2' osisID='Ps.7.2'/>Or else they will rip me apart</l>
     <l>like lions<lb type='x-secondLine'/>attacking a victim,<lb/>and no one
        will save me.<verse eID='Ps.7.2'/>
     </l>
   </lg>
</div>
```

A particular use of the **l** element is for marking the Hebrew term "Selah" which occurs in many Psalms. "Selah" (however it is translated or transliterated) should be marked up like this:

```
<l type="selah">Selah</l>
```

The **l** element's **type** attribute can have the following values:

- attribution
- doxology
- refrain
- selah

As with other attribute values, values beginning with "x-" are also permitted.

11.5.2.1. Elements allowed in l elements

The **l** element allows the following elements to occur within it:

• **a** • **abbr** • **chapter** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **mentioned** • **milestone** • **name** • **note** • **q** • **reference** • **seg** • **speaker** • **transChange** • **verse** • **w**

It is a milestoneable element. It also allows for mixed (text and element) content.

11.5.3. lb

The **lb** element, or "line break", is an empty element used to mark line breaks that are the result of linguistically or poetically significant structure as well as part of the typography and layout. For example, a line might be broken to fit into a narrow column.

The **lb** element is also used to mark where line breaks occurred in an important copy text, or where they should be placed in a text to be rendered.

Bible typesetting has a long tradition involving placement of such breaks. In some cases, translators have carefully decided preferred or required break-points for various set widths. These can be accommodated by using the **type** attribute of the **lb** element. For example,

```
type="x-wide-pref"
```

and

```
type="x-narrow-pref"
```

might be used to identify the locations of preferred line-breaks for wide and narrow column layouts. Similarly, type might be used to distinguish various levels of indentation following the break, or other typographic factors deemed important.

Note that the **type** attribute value for **lb** elements always has the "x-" extension in front of its value. The number of types of line breaks needed by different users varies from project to project and is no doubt always changing. Therefore the OSIS project made no attempt to collect even a representative set of types of line breaks and provided this mechanism to accomodate any reason for inserting a line break. Users are cautioned that they should be consistent in their use of this attribute on the **lb** element and it should be documented in the **description** element in the **header** of the document.

With the addition of the **editions** global attribute, it is also an option to define separate narrow and wide editions, and mark the preferred line breaks for each, indicating their applicability using the **editions** attribute.

The **lb** element should not be used merely to record where line breaks in general happened to occur in a source edition. For most source editions this information is unimportant; for ancient manuscripts it may be important, but must be marked up using the **milestone** element instead.

The **lb** element marks a point in the text and so have no content, either in the way of text or elements. It does have the attribute values described above but those are not considered to be element content.

11.6. Lists, tables, genealogies, figures and other material

Lists are very common in biblical and other texts. While one use mentioned below is for simple glossaries, lists appear and can be encoded using the **list** element wherever they appear.

Simple glossaries such as appear at the back of many Bibles, may be encoded at this time using the simple

list, **label**, **item** elements described below.

11.6.1. list

All types of lists are marked using the **list** element. OSIS does not define any types for lists but they can be distinguished by inserting a value for the **type** attribute. Recall that all undefined type values must begin with "x-". A **list** consists any number of **item** elements, preceded by optional **label** elements, which corresponded to the definition-terms of definition lists in various schemas.

The **list** element in OSIS is a good deal broader in application than the similar element found in TEI. It functions to organize any number of **item** elements but can also serve as a list of other containers.

11.6.1.1. Elements allowed in list elements

Elements that can occur in **list** are as follows:

• chapter • head • index • item • lb • list • milestone • q • verse

11.6.2. label

A leading **label** for a given **list item**. **Label** elements are optional. Encoders are cautioned, however, to be consistent in their use of **Label** elements as inconsistent usage may lead to unexpected results when formatting a text for print or display.

The **role** attribute is used to indicate the function of the **label** element in the list. Its value is unconstrained. Careful users will be consistent in the values they use for this attribute and document the values they use for it.

11.6.2.1. Elements allowed in label elements

The **label** element allows the following elements to occur within it:

• **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **milestone** • **name** • **note** • **q** • **reference** • **seg** • **transChange** • **w**

It is a milestoneable element. It also allows for mixed (text and element) content.

11.6.3. item

The main content of a list is encoded using the **item** element.

A **list** element can be used for outlines that sometimes preceed or follow a biblical passage, such as:

```
<list>
<head>Outline</head>
<item><label>I.</label> The Feasts of Xerxes (1:1-2.18)</item>
<item>
```

67

```
        <list>
<item><label>A.</label> Vashti Deposed (ch. 1)</item>
<item><label>B.</label> Esther Made Queen (2:1-18)</item>
        </list>
</item>
</list>
```

In the example the second **list** element is wholly contained within a **item** element. This corresponds to the normal TEI practice. The odds are against you ever needing any of the other elements allowed by **list** other than **item** but they are there should the need arise.

The **role** attribute on the **item** element is used to indicate a particular use of the item in the list. The value of that attribute is unconstrained. Users are cautioned to be consistent in the values used for it and to document that usage in the **description** element in the **header**.

11.6.3.1. Elements allowed in item elements

The **item** element allows the following elements to occur within it:

• **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lg** • **milestone** • **name** • **note** • **q** • **reference** • **seg** • **title** • **transChange** • **verse** • **w**

It is a milestoneable element. It also allows for mixed (text and element) content.

11.6.4. table

OSIS provides only very rudimentary tables: a table consists of rows, which in turn consist of cells. Formatting and layout is not part of the table markup; it can either be done automatically, as in HTML browsers, or by inserting some signal to the layout engine, such as type attributes or processing instructions. Note that a **table** can be nested inside another **table**. Simply start a new **table** element inside a **cell** element.

You can specify how many rows and columns there are with the attributes **rows** and **cols**, but this is optional. The value of either attribute must be a positive number. In the **head** element you can give the table a title or small description.

```
<table>
<row><cell role="label">Tribe </cell><cell role="label">Administrator</cell></row>
<row><cell>Reuben </cell><cell>Eliezer son of Zichri</cell></row>
<row><cell>Simeon </cell><cell>Shephatiah son of Maacah</cell></row>
<row><cell>Levi </cell><cell>Hashabiah son of Kemuel</cell></row>
<row><cell>Aaron </cell><cell>Zadok</cell></row>
<row><cell>Judah </cell><cell>Elihu, one of King David's brothers</cell></row>
...
</table>
```

The **type** attribute of the **row** element may have one of two values: label or data. Those values are used by processing software to create the proper presentation of the data in a row.

11.6.4.1. row

The **row** element binds together a number of cells for a line. The use is similar to the (X)HTML element

<tr>.

11.6.4.2. cell

Due to the fact that table cells in printed publications can contain nearly anything, nearly anything is allowed in OSIS **cell**s even verses and other tables.

The content of **cell** elements often needs alignment to be specified for proper rendering. To meet that need, the **cell** element has an **align** attribute that takes one of the following values:

- left
- right
- center
- justify
- start
- end

This example shows the use of **table**, **row** and **cell**. You see that in a cell you can use other elements from the OSIS scheme. Note in particular the **name** and **reference** elements.

```
<table cols="3" rows="4">
  <head>Settlement of the three patriarchs</head>
  <row>
    <cell>The group of</cell>
    <cell>settled nearby</cell>
    <cell>belonged to the tribe of</cell>
  </row>
  <row>
    <cell><name type="person">Abraham</name></cell>
    <cell><name type="geographic">Mamre</name>
        nearby
        <name type="geographic">Hebron</name>
        (eg. <reference osisRef="Gen.18">Gen 18</reference>,
        <reference osisRef="Gen.13.18">Gen 13:18</reference> or
        <reference osisRef="Gen.14.13">Gen 14:13</reference>. And
        more in the south:
        <reference osisRef="Gen.12.9">Gen 12:9</reference> or
        <reference osisRef="Gen.13.1">Gen 13:1</reference>)
    </cell>
    <cell><name type="x-group">Caleb</name> (
        <reference osisRef="Num.13">Num 13</reference>) and
later:
        <name type="x-group">Judah</name>).
    </cell>
  </row>
  <row>
    <cell><name type="person">Isaac</name></cell>
    <cell><name type="geographic">Beersheba</name>
        (<reference osisRef="Gen.26.23-Gen.26.33">Gen 26:23-33
         </reference>)
        and
        <name type="geographic">Beerlahairoi</name>
        in the South (<reference osisRef="Gen.24.62">Gen 24:62
         </reference>)
    </cell>
    <cell><name type="x-group">Simeon</name>
```

```
                in the South and
                <name type="x-group">Joseph</name>
                in the North
            </cell>
        </row>
        <row>
            <cell><name type="person">Jacob</name></cell>
            <cell><name
    type="geographic">Peni'el-Mahanaim-Succoth</name>,
                <name type="geographic">Sichem</name>
                and <name type="geographic">Bethel</name>
            </cell>
            <cell><name type="x-group">Reuben</name>, later
                <name type="x-group">Ephraim-Manasseh</name>
            </cell>
        </row>
    </table>
```

11.6.5. figure

The **figure** element is used to insert graphic non-textual materials, in other words, maps, pictures, drawings into an encoded text.

An example of a **figure** in an OSIS text might be:

```
<figure src="Beckmann_1917.jpg" alt="Painting by Max Beckmann, titled
        Christ and the Woman taken in Adultery"><caption>Christ and
        the Woman Taken in Adultery by Max Beckmann,
        1917</caption><index index="illustrations"
        index1="Beckmann, Max"/>
</figure>
```

At first it may look odd that the material in the **alt** attribute is repeated in the **caption** element. The **alt** attribute is important for situations where the application or user (for the visually impaired) cannot use or see the image that has been inserted in the text. The **alt** attribute is a friendly way of insuring that the encoded text will be understandable by the widest range of both applications and users.

The **src** attribute is used to point to the source for the image to be used in the text. It can be used to point to any type of image, including an SVG (Scaleable Vector Graphic) image. Future versions of the OSIS schema will include standard mechanisms for embedding SVG images directly in OSIS files.

Additional metadata about the image can be recorded in optional attributes for **catalog**, **location**, **osisRef**, **rights**, and **size**. Those attributes are used as follows:

- **catalog** A unique item or record number for the image.
- **location** The location depicted in an image, such as a town or area.
- **osisRef** Points to material that the figure depicts or illustrates.
- **rights** The holder of reproduction rights to the image or illustration.
- **size** Size of the image, useful for scaling for proper display.

The **index** element allows the encoder to encode the information necessary to automatically create an index, for either an online version of this material or a more traditional back of the book index. The **index** attribute gives the type of index where this item will appear, and **level1** provides the material that will appear in that index. See **index** (below) for more information on this element.

Another example for **figure** showing the use of the **osisRef** attribute.

```
<figure src="dore1.jpg"
  alt="Woodcut by G. Dore, titled The Agony in the Garden"
  osisRef="Luke.22.43">
  <caption>The Agony in the Garden by Gustave Dore,
1855</caption>
  <index index="illustrations" level1="woodcuts" level2="Dore,
Gustave" index3="Agony"/>
  <note type="allusion">DORE, Gustave, grafic artist, painter and
      sculptor, born 1/6/1832 in Strasburg, died 1/23/1883 in
      Paris. - For over 90 books of world literature he has
      created woodcutten illustrations and is especially known
      for his illustrations of the bible. This work here is an
      example for his extraordinary talent.</note>
</figure>
```

11.6.5.1. Elements allowed in figure elements

• **caption** • **index** • **note**

11.6.6. caption

The **caption** element is used with **figure** to provide the caption shown with a map, illustration or image.

11.6.6.1. Elements allowed in caption elements

The **caption** element allows the following elements to occur inside it:

• **a** • **abbr** • **divineName** • **foreign** • **hi** • **index** • **lb** • **milestone** • **name** • **note** • **q** • **reference** • **seg** • **w**

12. milestone

The **milestone** element is an empty element, and so is represented as a single element, thus **<milestone/>** rather than as a pair of elements. It is used to mark point events in a text, often involving the layout of the original text, or special points of access into the electronic text.

For example, when digitizing an ancient manuscript, it may be considered important to record where the page, column, and line boundaries of the original manuscript fell. This would be done as shown here:

```
<milestone type="pb" n="37-verso"/>
<p>The Lord said to Eliphaz:<milestone type="line"/>
What my servant Job has said about me is true, <milestone type="line"/>
but I am angry with you and your two friends for <milestone type="line"/>
not telling the truth. <verse osisID="Job.42.8">So I want you to go
over to <milestone type="line"/>
Job and offer seven bulls and seven goats on an <milestone type="line"/>
alter as a sacrifice to please me. After this, Job <milestone type="line"/>
will pray, and I will agree not to punush you for <milestone
type="line"/>your foolishness.</verse><milestone type="line"/>
<verse osisID="Job.42.9">Eliphaz, Bildad, and Zophar obeyed the Lord,
and he answered Job's prayer.</verse>....</p>
```

Note that because **milestone** is an empty or point element, not a container; it may be placed freely without concern about violating the boundaries of other elements in the same region.

Where a break to be represented by a **milestone** occurs between other units, such as verses or paragraphs, the **milestone** should be placed between those units, rather than just within either one.

When setting the attribute **n** on a **milestone**, it should indicate the number of the unit starting, not the unit ending. For example, **<milestone type="page" n="3"/>** indicates the break between pages 2 and 3, not between pages 3 and 4. Numbering does not need to be unique across various types of milestones -- for example, the 24th line on page 5 of a manuscript may be marked simply n="24", rather than n="5.24" or something similar.

Several predefined types are provided for the **milestone** element (the value for the **type** attribute is shown in bold):

- **pb**

  Marks the location of a page break in the source text.

- **column**

  Marks the location of a column break in the source text. Assuming page boundaries are also marked, the start of the first column need not be marked unless something else (such as a footer) precedes it in the encoding of the page. Columns should be numbered in the order of reading (for example, right to left in Hebrew texts). In the case of, say, an English/Hebrew diglot edition, where there is no principled order of reading among the columns, the direction used for the pages (Hebrew or Greek) should be considered the dominant direction, and the same direction should be used for numbering columns.

- **Cquote**

  Marks the location of a continuation quote mark in the text.

- **header**

  A milestone of type 'header' should precede the encoding of the page header if it is being included in the encoded text. This would normally be true only for digitized editions of manuscripts or other important copy editions, because in modern print Bibles headers are typically automatically generated.

- **footer**

  Type 'footer' should be used just like type 'header,' except that it marks the page footer area instead.

- **line**

  Line milestones should be used to mark line breaks in the copy text when they are considered significant. This will normally only be true for important manuscripts, where line numbering may be needed for paleographic or reference use. Line milestones must not be used to represent linguistically significant line breaks, such as in poetry, for which the lg and l elements are provided.

- **halfLine**

  In certain languages it is important to mark half-line units, and this type is provided for such cases.

- **screen**

  The milestone of type 'screen' is to be used to mark preferred break points in an on-screen rendering of the text. For example, if the user requests to be taken to the book of Psalms in a given electronic edition, it may be best not to take them to Psalm.1.1, but to an earlier point, preceding any introductory material. In many cases this can be accomplished by taking them to the appropriate div (since the **<div type="book" osisID="Ps">** should precede and Psalms-specific introductory material); but this milestone type is available for other cases. The OSIS specification does not impose requirements on how applications make use of such milestones.

The **milestone** element also has a **marker** attribute. This attribute is used to record typographic features in a text that cannot be duplicated by automated scripting. If a text inconsistently using a paragraph marker for example, it would not be possible to duplicate inconsistent usage without that information being recorded as part of the encoding process. Those cases are thought to be rare and should you discover one, please send a note to osis-editors@bibletechnologieswg.org.

13. Common elements in all texts

The elements found in this section can be found in almost any encoded text.

13.1. a

The **a** element is similar to the HTML **a** element, and likewise may be used to encode links within a document. This eases integration of OSIS documents into the Web environment. For example:

```
    <p>See Edwards' famous treatise on <a
href="http://www.ccel.org/e/edwards/affections/religious_affections.html">religious
affections</a> for additional information.</p>
```

The **a** element has all the characteristics of the **a** element in HTML. This element can be used with any defined pointer system from the W3C to point into multimedia or other files, as allowed by some defined reference system.

The **href** attribute is required on all **a** elements and the value must conform to W3C requirements.

a element allows a mixture of text and the following elements to occur within it:

- **index**

13.2. abbr

The **abbr** element marks a portion of text as an abbreviation.

While an abbreviation may be familiar to you, it may be unknown to many readers. It is not required but strongly suggested that every abbreviation in a text should be encoded and the **expansion** attribute used to record what the abbreviation represents. For example:

```
<abbr expansion="Journal of Biblical Literature">JBL</abbr>
```

This element will be most often used in notes, where citations are abbreviated and users may not be familiar with the abbreviation. Putting the expansion of the abbreviation in the **expansion** attribute allows software to choose to display the expanded form instead of the abbreviation or to display it upon request by the reader.

13.2.1. Abbreviation Examples

8:7 *Or on the pole.*   **8:9a** Hebrew *on the twenty-third day of the third month, the month of Sivan,* of the Hebrew calendar. This event occurred on June 25, 474 ʙᴄ; also see note on 2:16.   **8:9b** Hebrew *to Cush.* **8:12** Hebrew *the thirteenth day of the twelfth month, the month of Adar,* of the Hebrew calendar. The date selected was March 7, 473 ʙᴄ; also see note on 2:16.   **9:1** Hebrew *on the thirteenth day of the twelfth month, the month of Adar,* of the Hebrew calendar. This event occurred on March 7, 473 ʙᴄ; also see note on 2:16.

The highlighted text, in this case, **BC** is an abbreviation for **Before Christ**. That may seem too 'well known' to encode, but how many readers would recognize **BWL**? (hint, usually cited in studies on Job) In all but the most common cases, abbreviations should be marked and the expanded forms included.

> David" (KJV). "Then David rested with his fathers and was buried in the City of David" (NIV). "Then David died and was buried in the City of David" (NLT). Only the New Living Translation clearly translates the real meaning of the Hebrew idiom "slept with his fathers" into contemporary English.

The highlighted text, in this case, **NIV**, **KJV**, and, **NLT** are abbreviations for **New International Version**, **King James Version**, and **New Living Translation**, respectively.

13.2.2. Elements allowed in abbr elements

abbr element allows the following elements to occur within it:

• **a** • **divineName** • **foreign** • **index** • **name** • **note** • **reference** • **w**

It is a milestoneable element. It also allows for mixed (text and element) content.

13.3. catchWord

Catchwords and catchphrases are those parts of notes that are copied from the main text, to orient the reader as to the note's precise applicability. Catchwords in notes must be marked when present. For example:

```
<verse osisID="NRSV:Ezek.19.5">When she saw that she was thwarted,
that her hope was lost, she took another of her cubs and made him a
young lion.</verse> <note>It is uncertain to which king <catchWord
osisRef="Ezek.19.5">another of her cubs</catchWord> refers....</note>
```

Users of XML markup software will note that **catchWord** is allowed in the **p** element. That is only because there can be paragraphs inside of notes. OSIS conformant documents must limit their use of **catchWord** to being within **note** elements only.

The **osisRef** attribute on this element can be used to point to the exact word or set of words that are acting as the catchword or phrase. This will be useful when searching groups of texts for all the catchwords or phrases for particular terms. Use of this attribute is optional.

13.3.1. catchWord Examples

> **Moses . . . well**   Read about this in Leviticus 14:1-32.

The bold text is encoded with the <catchWord> element to show that it appears in the main text. Only the first and last word of the phrase are shown.

> ᶜ **we need**; or for today, or for tomorrow.

The bold text here is also encoded with a <catchWord> element. The following **for today**, **for tomorrow**

should be encoded with the <rdg> element.

13.3.2. Elements allowed in catchWord elements

The **catchWord** element allows the following elements to occur within it:

• **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **inscription** • **lb** • **list** • **mentioned** • **milestone** • **name** • **note** • **q** • **reference** • **seg** • **speaker** • **title** • **transChange** • **w**

It is not a milestoneable element. It does allow for mixed (text and element) content.

13.4. divineName

The **divineName** element is used only for the Deity. Angels, demons, idols, and the like should be tagged with <name type='nonhuman'> For example:

```
<divineName>El Shaddai</divineName>
```

13.4.1. divineName Examples

The Garden of Eden

When the LORD*f* God made the universe, 5 there were no plants on the earth and no seeds had sprouted, because he had not sent any rain, and there was no one to cultivate the land; 6 but water would come up from beneath the surface and water the ground. 7 Then the LORD God took some soil from the ground*g* and formed a man*g* out of it; he breathed life-giving breath into his nostrils and the man began to live. 8 Then the LORD God planted a garden in Eden, in the East, and there he put the man he had formed. 9 He made all kinds

The divine name is rendered differently in some translations when the Tetragrammaton is encountered as opposed to Adonai.

> 13 But Moses replied, "When I go to the Israelites and say to them, 'The God of your ancestors sent me to you,' they will ask me, 'What is his name?' So what can I tell them?"
>
> 14 God said, "I am who I am. You must tell them: 'The one who is called I AM/ has sent me to you.' 15 Tell the Israelites that I, the LORD, the God of their ancestors, the God of Abraham, Isaac, and Jacob, have sent you to them. This is my name forever; this is what all future generations are to call me. 16 Go and gather

Another example of divine name.

13.4.2. Elements allowed in divineName elements

The **divineName** element allows the following elements to occur within it:

• **a** • **abbr** • **foreign** • **index** • **note** • **reference** • **seg** • **w**

It is not a milestoneable element. It does allow for mixed (text and element) content.

13.5. foreign

The **foreign** element is used to mark a portion of text that is not in the primary language of the text, such as 'Talitha Cum' in Mark 5:41. The specific language should be indicated via the xml:lang attribute. For example:

```
<verse osisID="NRSV:Mark.5.41">He took her by the hand and said to her:
<q><foreign xml:lang="arc">Talitha cum</foreign></q>, which means,
<q>Little girl, get up!</q></verse>
```

13.5.1. foreign Examples

The Death of Jesus
(Mark 15.33-41; Luke 23.44-49; John 19.28-30)

> 45 At noon the whole country was covered with darkness, which lasted for three hours. 46 At about three o'clock Jesus cried out with a loud shout, "Eli, Eli, lema sabachthani?" which means, "My God, my God, why did you abandon me?"

An example of a foreign phrase that is retained in a translation. Note that the foreign phrase is contained in a quotation so both <q> and <foreign> elements would be used to mark this phrase.

their house send warm Christian greet-ings. ²⁰ All the believers here send greetings.

Greet one another with the kiss of peace.

²¹ With my own hand I write this: *Greetings from Paul.*

²² Whoever does not love the Lord — a curse on him!

Marana tha — Our Lord, come!

²³ The grace of the Lord Jesus be with you.

²⁴ My love be with you all in Christ Jesus.

Another example of a foreign phrase.

13.5.2. Elements allowed in foreign elements

The **foreign** element allows the following elements to occur within it:

• **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **milestone** • **name** • **note** • **reference** • **seg** • **title** • **w**

It is a milestoneable element. It also allows for mixed (text and element) content.

13.6. hi

The **hi** element provides a simple text highlighting mechanism. If the highlighted portion of text is marking a foreign word, a title, or anything covered by a specific OSIS element, then the appropriate OSIS element should be used. In the two cases cited, a foreign word would be encoded with the **foreign** element and a header would be encoded with the **title** element. The more specific encoding of words in a text allows for better searching and rendition of the text than simply highlighting the text. The **hi** element is reserved for cases where the reason for the highlighting or other typographic distinction is unknown.

The type attribute on the **hi** element allows the user to specify what typographic distinction was observed in the text. As noted above this is not meant as a guide for stylesheets but for recording what was observed. If it is known with reasonable certainty why a word or phrase appears in italic, for example a foreign phrase, then the **foreign** element should be use to mark it. To promote consistency in the marking of such distinct texts, the OSIS schema provides eleven (11) standard values for the type attribute on the **hi** element as follows:

- **acrostic**
- **bold**
- **emphasis**
- **illuminated**
- **italic**
- **line-through**
- **normal**

- **small-caps**
- **sub**
- **super**
- **underline**

If additional values are needed, they should be created by prepending "x-" to the value.

```
<catchWord><hi type="italic">the child with his mother Mary
    </hi>
    </catchword>
```

It should be noted that Lord in small-caps could be encoded using **divineName** without the use of the **hi** element.

13.6.1. Elements allowed in hi elements

hi element allows the following elements to occur within it:

• **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **milestone** • **name** • **note** • **reference** • **seg** • **w**

It is not a milestoneable element. It does allow for mixed (text and element) content.

13.7. index

The **index** element may be placed at any point in the document to indicate a topic under which that location should be indexed. It is always an empty element. Multiple indexes (such as of places, names, theological or ethical issues, etc) must be distinguished via the **name** attribute.

The particular index in which an entry should appear is given by the **index** attribute of this element.

Indexes with up to 4 levels of headings are supported. The primary index entry name is specified on the **level1** attribute, followed by sub-headings **level2**, **level3**, and **level4**. For example:

```
<title>On Justice<index index="topic" level1="Virtues"
level2="Justice"/></title>
```

Note that the use of **level1** is required if an **index** element is used, and that use of **level3** requires the presence of **level2**, just as the use of **level4** requires the presence of **level3** attributes.

There is also a **see** attribute, which may be used to represent the need for a cross-reference to another **index** entry; such elements should be placed together at the end of the document body (since they do not refer to a particular location). For example:

```
<index index="topic" level1="Virtues" level2="Justice" see="Fairness"/>
```

The **index** element represents a point in the text and therefore does not allow any content, either text or elements.

13.8. inscription

The **inscription** element is used to encode inscriptions in a text. Such inscriptions should not also be tagged as quotations. For example, where Paul refers to an alter inscription in Athens (NIV):

```
<verse osisID="Acts.17.23">For as I walked around and looked carefully
    at your objects of worship, I even found an altar with this inscription:
    <inscription>To an unknown god</inscription>
```

Another example is the inscription in the book of Daniel:

```
<p><verse osisID="Dan.5.22"><q sID="q.Dan.83"/ marker="'">But you
his son, O Belshazzar, have not humbled yourself, though you knew all
this.</verse>
<verse osisID="Dan.5.23">Instead, you have set yourself up against
the Lord of heaven. You had the goblets from his temple brought to
you, and you and your nobles, your wives and your concubines drank
wine from them. You praised the gods of silver and gold, of bronze,
iron, wood and stone, which cannot see or hear or understand. But you
did not honor the God who holds in his hand your life and all your
ways.</verse>
<verse osisID="Dan.5.24">Therefore he sent the hand that wrote the
inscription.</verse></p>
<p><verse osisID="Dan.5.25">This is the inscription that was written:
<foreign xml:lang="arc"><inscription>Mene, Mene, Tekel,
Parsin</inscription><foreign></verse>
```

A milestone form of the **q** element is used to represent the start of Daniel's quotation, which ends at the end of verse 28 (not shown).

In the example from Daniel, the repetition of words from the inscription (in verse 26-28) should not be marked as inscriptions.

Inscriptions are found in Exod.39.30, Dan.5.25, 2Tim.2.19. There are additional passages where inscriptions are mentioned without being quoted verbatim, such as Matt.22.20; these would not be encoded using the **inscription** element.

13.8.1. Elements allowed in inscription elements

The **inscription** element allows the following elements to occur within it:

• **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **milestone** • **name** • **note** • **q** • **reference** • **seg** • **speaker** • **speech** • **transChange** • **w**

It is not a milestoneable element. It allows for mixed (text and element) content.

13.9. mentioned

The **mentioned** element marks meta-linguistic use of a term. That is, it encloses a word, phrase, or other unit that is not being *used*, but only *mentioned*. For example:

```
<verse osisID="NRSV:John.1.42">He brought Simon to Jesus, who looked
at him and said, <q>Your are Simon son of John. You are to be called
<mentioned>Cephas</mentioned></q> (which is translated Peter).</verse>
```

In this example, *Cephas* is not being used by Jesus to call Simon to him but is being used to tell him his new name.

13.9.1. Elements allowed in mentioned elements

The **mentioned** element allows the following elements to occur within it:

• **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **milestone** • **name** • **note** • **q** • **reference** • **seg** • **transChange** • **w**

It is not a milestoneable element. It allows for mixed (text and element) content.

13.10. name

When a *name* appears in a text, it is important to mark it with this element and to use the type attribute to record what type of name has been marked. Remember that a computer cannot distinguish *Job,* as in the man from Ur, from *job,* as in 'I have a job for you...' without your assistance, at least at the beginning of a sentence. Despite what you may have read in the news media, computers are very literal and quite dumb when it comes to reading texts.

There are any number of types of names and OSIS has specified a small set of the most useful ones for use when encoding a Bible. If others are needed, remember to start the value with "x-" to distinguish it from those defined by OSIS.

- **geographic**
- **holiday**
- **nonhuman**
- **person**
- **ritual**

```
<verse osisID="NRSV:Job.1.1">There once was a man in the land of
<name type="geographic">Uz</name> whose name was
<name type="person">Job</name>. That man was blameless and upright, one who feared
<divineName>God</divineName> and turned away from evil.</verse>
```

Note that there are three names in that verse, one geographic, one of a person, and one of the Deity. The first two are marked with the *name* element and appropriate type attribute. Any use of any form of the name of the Deity is marked with *divineName* Different forms of the divine name should be specified, if desired, using the **type** attribute on **divineName**.

It may be the case that the same person or place is referencd in a text by different names or different forms of the same name. For example, we all know that if the **name** element was used to mark *Jesus* and *Jesus of Nazareth*, that it is the same person. A computer on the other hand, would not. In such cases, use the **regular** attribute on both, **regular = "Jesus"** and automated searches will return both cases.

13.10.1. Elements allowed in name elements

The **name** element allows the following elements to occur within it:

• **a** • **abbr** • **foreign** • **index** • **note** • **reference** • **seg** • **w**

It is not a milestoneable element. It allows for mixed (text and element) content.

13.11. rdg

Bibles and other ancient works frequently differ in different surviving manuscripts. Sometimes this variation must be represented in printed, electronic, or other editions, even in translation. For such cases, OSIS provides the **rdg** ("reading") element, corresponding to the TEI element of the same name. This element is used to identify each variation of a particular portion of the text, and may identify each variant with the set of textual witnesses that support it and with a type such as primary, secondary, etc.

The **rdg** element contains the variation from the main text and appears in a **note** element as follows:

```
<verse osisID="NRSV.Song.2.1">I am a rose
     <note osisRef="NRSV.Song.2.1@s[rose]">Heb
         <rdg>crocus</rdg>
     </note> of Sharon,
a lily of the valleys.</verse>
```

This example illustrates (or reinforces) several points: 1. A note appears directly in the textual material where the user would normally see a raised letter or number to indicate a note. 2. The **osisRef** attribute allows the note to point at a particular word in the text to which the note applies. 3. The **rdg** element holds an alternative word or reading to the one found in the text. The interested reader will note that the identification of "crocus" is unclear but it is known that there were no "roses" in the modern sense of the word growing on the plain of Sharon (northern Israel) in biblical times.

The **rdg** element has several pre-defined value for type:

- **alternate**: A reading of approximately equal probability compared to others.
- **variant**: A reading that varies from the accepted tradition.

If the enumerated values for type on the **rdg** element are insufficient, users should create their own, but prepended with the 'x-' extension.

The **witness** attribute should be used with a value conforming to the **osisRef** syntax to indicate the source of the reading.

13.11.1. Elements allowed in rdg elements

The **rdg** element allows the following elements to occur within it:

• **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **lg** • **list** • **milestone** • **name** • **note** • **p** • **q** • **reference** • **seg** • **transChange** • **w**

It is not a milestoneable element. It does allow for mixed (text and element) content.

13.12. rdgGroup

The **rdgGroup** element is used only to group together **rdg** elements. It will be used most often to collect readings that share a common source or tradition.

13.13. reference

The **reference** element is used to encode an explicit cross-reference to another passage or work (the work referred to need not be Biblical, but must be declared via a **work** element in the header, and by accessible via the same canonical referencing scheme defined in osisID syntax). Reference elements will often occur within notes, but may also occur freely in text (the latter is more common when encoding non-Biblical works).

The **type** attribute on this element has three enumerated values and can also have values supplied by the user which start with "x-." The OSIS defined values are:

- annotateRef: Meaning that this reference is an annotation of the referenced material.
- parallel: The referenced material is a parallel to the material under discussion.
- source: The referenced material is the source of the material under discussion.

The predefined values enable the encoder to record if the reference is meant to be an annotation on another work (annotateRef), or is a parallel to another work, such as with the Synoptic Gospels (Matthew, Mark, and Luke), or to point another work as the source of the material under discussion (such as an electronic version of a manuscript). Different interfaces might want to treat those references differently and hence the ability to record the type of the reference that is being made.

The **osisRef** attribute is used to specify in a machine processable way the target of the reference. A reader can easily use biblical citations written in any number of ways after a moments hesitation. Computers on the other hand will never find the correct material unless assisted by properly written references. OSIS provides one standard way to write such references.

A reference element was used in the **note** example above. To refresh your memory, here is just the **reference** element part of that example:

```
<reference osisRef="Ezra.4.6">Ezra 4:6<reference>
```

Note there is no **osisWork** prefix, that is no 'name:' in from of 'Ezra.4.6' in the **osisRef** attribute. That may be for one of two reasons: First, that is being supplied by the **osisWork** default, i.e., it is a reference in this work. Or, the **osisWork** prefix may have been set by the the **workPrefix** element in the **header** element of the document.

In either case, if you want to point to another text, you must declare that in a **work** element and use the value of the **osisWork** attribute from it to make references to it. For example, let's assume that the text has a **work** element declared for David Clines' commentary on Job. It has an **osisWork** attribute value of 'dclines_job.' Here is how to use that value to make a reference to that work:

```
<reference osisRef="dclines_job">Clines, <title>Job 1-20</title>, at p. 28<reference>
```

13.13.1. Elements allowed in reference elements

reference element allows the following elements to occur within it:

• **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **milestone** • **name** • **note** • **seg** • **title** • **w**

13.14. seg

The **seg** element used to mark small portions of text in ways not provided for by the OSIS schema. For example, the lowest level of division that has a defined element in OSIS is word.

In cases where subdivisions of words need to be encoded, prefixes, suffixes, morphemes, the **seg** element is the correct element to be used. It can also be used, with caution, to mark a textual feature that is not otherwise provided for by the schema. It should be noted that this element can only contain very small portions of text and cannot contain things like verses or paragraphs.

The **type** attribute on the **seg** element has four values (in addition to "x-") which are:

- alluded
- keyword
- otPassage
- verseNumber

```
<word><seg>W:</seg><seg>R74W.XA</seg></word>
```

Don't worry if you don't recognize the word or its notation. This was taken from the MORPH database produced by the Westminster Hebrew Institute.

13.14.1. Elements allowed in seg elements

The **seg** element allows the following elements to occur within it:

• **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **milestone** • **name** • **note** • **q** • **reference** • **seg** • **transChange** • **w**

It is a milestoneable element. It also allows for mixed (text and element) content.

13.15. transChange

The **transChange** element is used to mark text that was changed in a notable way in translation. For example, the KJV traditionally distinguishes all words inserted in translation (often by italics); the Amplified Bible has several punctuation conventions for marking explanatory or other expansions; and some translations indicate where the tense of verbs has been changed, perhaps due to sequence-of-tense requirements in the target language.

The **type** attribute is used to provide the type of change that has been made. OSIS has predefined values for this attribute, which can be used along with user defined values. User defined values must begin with "x-."

- **added**
- **amplified**
- **changed**
- **deleted**
- **implied**
- **moved**
- **tenseChange**

For example, the Hebrew text does not contain *woman* in the following example, so using **transChange** to mark the text:

```
<div osisID="NRSV.Song.2">
<speech>
        <transChange type="added">woman</transChange>
<verse osisID="NRSV.Song.2.1">I am a rose of Sharon, a lily of the valleys.</verse>
</speech>
</div>
```

13.15.1. Elements allowed in transChange elements

The **transChange** element allows the following elements to occur within it:

• **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **milestone** • **name** • **note** • **q** • **reference** • **seg** • **w**

It is not a milestoneable element. It does allow for mixed (text and element) content.

13.16. Verse

The history of verses in the Bible is an interesting one but full treatment of that history would take far more space than can be spared in this manual. Suffice it to say that the original texts of the Bible had no versification as is currently represented in various Bible editions. The translation of the Bible into English was not without controversy over the issue of versification. Current translation practice is moving towards presentation of the text as written and is less focused on the rather late occurrence of versification.

However, many important historical and contemporary editions of the Bible rely upon versification as both a means of organizing the text as well as for quick reference. Verses form an important part of the OSIS schema but it should be noted that in all but rare cases, the milestone form of the **verse** element should be used.

13.16.1. Elements allowed in verse elements

The **verse** element allows the following elements to occur within it:

• **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **inscription** • **lb** • **list** • **mentioned** • **milestone** • **name** • **note** • **q** • **reference** • **seg** • **speaker** • **title** • **transChange** • **w**

85

It is a milestoneable element. It does allow for mixed (text and element) content.

13.17. w

The **w** element provides a place to record word-level annotation, such as part of speech identifiers, lemma or Strong's numbers, and the like.

w element has the following attributes in addition to those that it shares with other elements:

- **gloss** Record comments on a particular word or its usage.
- **lemma** Use to record the base form of a word.
- **morph** Use to record grammatical information for a word.
- **POS** Use to record the function of a word according to a particular view of the language's syntax.
- **src** Use to record origin of the word.
- **xlit** Use to record a transliteration of a word.

In the following example, it is assumed that Strong's Numbers have been recorded in a **work** element and the "s" prefix seen on the **gloss** attribute is a reference to it. Recall that work prefixes can also be defaulted by use of the **workPrefix** element.

```
<word gloss="s:H325>Ahasuerus</word>
```

13.17.1. Elements allowed in w elements

The **w** element allows the following elements to occur within it:

• **a** • **index** • **note** • **seg**

It is not a milestoneable element. It does allow for mixed (text and element) content.

14. Bible Editions

Works are sometimes also published in multiple editions with slightly differing content. For example, Bibles may be published in Catholic and Protestant editions, which differ in the selection of Deutero-canonical books and possibly in other ways as well. Such differences should be expressed using the **editions** attribute (added in OSIS version 2.1), which is available on all OSIS elements.

When using the **editions** attribute, its value should contain one or more **osisWorkID**s, each associated with a specific edition declared by a **work** element in the **header**. These are the editions for which the content of the element is to be included. Any elements that are not in the scope of a non-null **editions** attribute, default to being included in all editions.

For example, to indicate that the book of Sirach is to be included in the "Catholic" and "Study" editions, but not others, the book would begin:

```
<div type="book" editions="Catholic Study" osisID="Sir">
   <title>Sirach</title>
```

15. osisIDs

A canonical reference scheme is a system of agreed names and/or numbers for referring to parts of a document. In the Bible, the traditional system used in most languages consists of a book name (such as Genesis), then a chapter number, then a verse number. Most works of Classical literature have similar schemes, nearly all of which are also hierarchical (that is, they work from larger units to smaller).

The basic form for Biblical verse references is strictly defined by OSIS, so that various electronic Bible versions can interoperate easily. Standard abbreviations for the canonical and deuterocanonical books are provided; chapter and verse numbers follow the book abbreviation separated by periods. For example:

```
Matt.1.1
```

OSIS uses such identifiers in several places:

- To identify a portion of text from an actual canonical work, such as a verse of the Bible. The verse element bears an **osisID** attribute which must include the identifier appropriate to the verse. For example, <verse osisID="Matt.1.1"/>....
- To identify a reference **to** a Biblical or other passage, that is not contained at the point of reference. For example, "<p>The correctness of my exegesis is incontrovertibly proven by <reference osisRef="Matt.1.1">the first verse of Matthew.</p>"
- In the header, to identify what portions of the Bible are included in a declared work. For example, a particular edition may include only the NT and Psalms. The **scope** element may be used to specify each relevant portion.

15.1. Partial identifiers

It is permissible to refer to an entire chapter by simply omitting the verse number and the preceding ".", for example:

```
Matt.5
```

Similarly, it is permissible to refer to an entire book by omitting the chapter and verse number and both corresponding periods:

```
1Cor
```

For those books of the Bible that have only 1 chapter, the chapter number "1" must be specified: The first verse of Jude is thus Jude.1.1, not Jude.1.

15.2. Works

References are used to identify a place in a particular edition or translation of the Bible, or to other works entirely, such as Josephus, writing of the Apostolic Fathers, classical or modern literature, and so on. Once a work has been declared in a **work** element, it has a short name that can be used with any reference to it, for example:

```
NIV:Matt.1.1
```

In this case, the **NIV:** is the short name that is declared in a **work** element in the **header** of this particular OSIS document. The colon is required to make it is clear where the work short name ends and the reference in the work begins.

It is possible to refer to an entire work, such as the whole CEV, NIV, KJV, the Iliad, etc. However, to do so the work name must be stated, and the following colon must be included (without the colon, it would be interpreted as a top-level identifier within the default work).

15.3. Sub-identifiers

Translations also often split verses into parts, provided labels such as 'a' and 'b' for the separate parts. Encoders may freely add sub-identifiers below the lowest standardized level. They are set off from the standardized portion by the character '!'. For example:

```
Rev.2.20!b
```

Such subdivisions are not standard across different translations, so applications must be prepared to discard them when trying to locate a referenced location in a different edition.

These extensions are not considered a formal part of the canonical reference scheme, and so a work that uses them need not claim it is using a different scheme.

15.4. Grouping

Translators often group several adjacent verses into a single block, so that they can translate them using word order more natural in the target language. In such cases, the larger unit (commonly a paragraph or **p** element), gets an **osisID** that lists all the individual **osisID**s for the verses included, separated by white space. For example:

```
<p osisID="Matt.1.1 Matt.1.2 Matt.1.3">...</p>
```

osisIDs never allow the use of ranges. Only **osisRef**s (discussed later) do.

Ranges are prohibited for **osisID**s in order to simplify implementation of tools that search for particular passages by reference. If an encoder wished to mark IDs at, say, the pericope level, the markup would be quite verbose because many verses would need to be listed in a single attribute on the **div** type="x-pericope". However, there is no need to do this if the verses within the pericope are themselves identified. (A "pericope" is a short portion or passage of Scripture, usually read in public worship.)

15.5. Other details of osisIDs

The "."-separated parts of an **osisID** are defined to represent a hierarchy. In the traditional versification (introduced by William Whittingham in his New Testament published in Geneva in 1557), these would be book, chapter, and verse numbers. In other schemes for the Bible, or schemes for entirely different works, the names of the parts may differ, but the expectation is that they still form a hierarchy.

89

The parts of an **osisID** may contain any mixture of numbers, letters, and underscores. However, to avoid conflict with the other punctuations used (such as ":" to separate the work from the in-work location, "@" to separate fine-grained references in osisRefs, and "!" to separate work-specific extensions to a reference system), no other characters are allowed.

The **!** as the terminator, allows encoders to append names and/or numbers to provide finer-grained reference points. Such extensions may not be valid across reference systems so should be used with care.

As with XML in general, these identifiers are case-sensitive: "Matt" is the correct form, and there is no Bible book in OSIS that is called "matt" or "MATT". Applications for end users may choose to accept case variants in such names, but applications for encoders (such as OSIS editors) must not produce documents with invalid reference names.

Most users will not expect to see **Matt.1.1** or similar forms of verse references when reading a Bible. The format of **osisID**s and **osisRef**s were designed for computer and not human consumption. The display form of either an **osisID** or **osisRef** (or any other attribute value for that matter) is limited only by the imagination of the person rendering the OSIS encoded text for reading. Thus, **Matt.1.1**, can be rendered as: **Matthew 1:1**, or **Matt. 1:1**, or **Matthew 1.1** or read aloud (for visually imparied users), or any other rendering that is desired. The 'encoded' form of any **osisID** or **osisRef** is only the starting point for rendering.

16. osisRefs

An **osisRef** is very much like an **osisID**. The fundamental difference is that while an **osisID identifies** the actual occurrence of canonical text, an **osisRef** is used to **refer** to canonical text from somewhere else. For example, a footnote (particularly one of type="crossReference") may refer to a related passage, or a section heading in Mark may include references to the parallel passages in Matthew and Luke; in such cases an **osisRef** rather than an **osisID** is used.

Any valid **osisID** value is also a valid **osisRef** value, and refers to the same thing. Thus for example, a commentary might say:

```
<p>The same interpretive method applies also in <reference
osisRef="Luke.1.1">the first verse of Luke</reference>.</p>
```

However, **osisRef**s provide additional capabilities. They can refer to a contiguous range of books, chapters, verses (or other units, as applicable to the work being referenced), and they can refer to precise locations within a given canonically-reference unit.

To refer to a range, simply include two **osisRef**s, one for the first verse (or chapter or book) of the range, and one for the last. Separate the 2 values by a single hyphen. For example:

```
John.3.14-John.3.16
Prov.30-Prov.31
Esth-Song
Ps.149-Prov.3.4
```

Both sides of the hyphen must hold complete references. It is not correct to abbreviate the first example above to merely "John.3.14-16" (as always, the values of **osisID** and **osisRef** attributes need not be the same values displayed to the reader).

A single **osisRef** cannot identify a discontiguous range of a work. For example, a complex reference such as "John 3:14-16, 18; 4:1-2; 19-20" cannot be encoded as a single reference. It must instead be encoded as several parts, each contiguous:

```
<p>See also
    <reference osisRef="John.3.14-John.3.16">John 3:14-16, </reference>
    <reference osisRef="John.3.18">18; </reference>
    <reference osisRef="John.4.1-John.4.2">4:1-2; </reference>
    <reference osisRef="John.4.19-John.4.20">4:19-20; </reference>.</p>
```

It is permissible for **osisRef** values, including those on either side of a hyphen in a range reference, to use **osisID** values that include the work-specific extension fields ("!" followed by a name, e.g. osisRef="Ps.3.23!b-Ps.3.24!a")).

16.1. Fine Grained References

To refer to specific locations within a named canonical reference element, give the **osisID** as usual, followed by a "grain identifier", which consists of the character "@", and then an identifier for the portion desired. Such identifiers are of the form:

```
grainType(parameters)
```

Two grain types are defined at this time:

The **cp** operator (short for "code point") counts through the character content of the referenced element, essentially by characters. Technically, the units counted are Unicode code points, a term which is defined more precisely than the generic term "character". The first code point of content is number 1, not 0.

Markup does not imply a space for purposes of counting even if it may for purposes of layout, printing or indexing. In other words, when counting the spaces for the **cp** operator, the user does not include any markup or its content.

To address a range of code points, the operator must be written as a range, thus:

```
Matt.1.1@cp(3)-Matt.1.1@cp(12)
```

When referring to a location in canonical content, text within non-canonical elements is not counted. (Thus, the intuitive count will not be changed by the insertion of notes, references, critical apparatus, and the like). When referring to a location in non-canonical content, all text in all included elements counts, whether canonical or not.

The **s** operator (short for string) finds the first match of the string value specified, with regard to case, within the canonical reference specified. If the canonical reference is one of several applied to the same target element (for example, when a paragraph has osisRef="Matt.1.1 Matt.1.2 Matt.1.3), that whole element is searched. If the string is not found, the user agent must warn the user, and may offer to suggest a best guess (for example, by searching again while ignoring case, whitespace, punctuation, accents, spelling variation, etc).

17. Different versification systems

Hebrew tradition numbers the proscriptions above Psalms (such as "A Maskil according to David") as verse one, and goes on from there; Greek tradition does not number the proscriptions, and start verse 1 after it. Of course, this would make all references in Psalms be off by one verse if the version reached is from the other tradition.

A few languages use traditional reference schemes that completely differ from the familiar book/chapter/verse one.

Many works of Classical literature likewise have more than one standard canonical reference scheme, such a Loeb numbers plus another method.

In such cases, where there are large and systematic differences, different "versification schemes" must be defined and named. On the other hand, nearly every edition of the Bible has some slight deviations from a standard versification scheme that it otherwise follows: for example, subdividing verses into parts "a" and "b", combining verses into a larger translation unit, and so on. It is highly undesirable to call these separate versification schemes, because they differ so slightly; because the differences can be mechanically resolved; and because there is considerably overhead to maintaining and mapping among versification schemes. Thus, as described below such minor extensions can be done without an edition having to say it is using a completely different versification scheme.

Each work must identify which versification scheme(s) it uses; this is done by a reference to the versification scheme declared by a **work** declaration in the header except that the predefined versification systems need not have work declarations.

References can also state what versification scheme they are expressed in, so that they may be correctly interpreted.

18. Conclusion

The OSIS markup documented in this users guide is in active use in publishing and translation projects. The needs of such projects will change as new mediums for delivery of the biblical text arise and Bibles are produced to meet the needs of diverse audiences. OSIS markup will change as well to meet the demands of such projects.

However, every effort will be made to insure that users of OSIS markup always have an easy entrance ramp as it were so they can gain the benefits of using markup to facilitate engagement with the Bible. You can play a role in that effort by joining in discussions of new needs or requirements for publishing, delivery and translating the Bible and what OSIS needs to do to meet those needs.

Please join the OSIS Users' Group. To do so, send mail to osis-user@whi.wts.edu, setting the Subject line to "subscribe". Online information about OSIS is also available at http://www.bibletechnologies.org and http://www.bibletechnologieswg.org.

The theme or motto of OSIS has been 'A Common Format for Many Visions.' It will have succeeded only when you are able to express your vision of the Bible using OSIS.

Appendix A OSIS Elements

Fuller explanations of all these elements are provided in the main body of this manual. However, after becoming familiar with OSIS, users may find it more useful to have a short listing of elements with a minimum of explanation for each one. The following is a "cheat sheet" as it were to OSIS. Note that in the event of uncertainty, the fuller explanation in the manual should be consulted.

- **a:** Similar to an HTML link. The "http://..." is recorded in the **href** attribute of this element. The "http://..." should not appear in the content of this element but rather the name or other description of of the link.

  In addition to text, it may contain **index** elements.

  Attributes of the <a> element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **href** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subtype** • **xml:lang** • **xml:space**

- **abbr:** Abbreviations should be encoded using this element. The **expansion** attribute records the full form of what was abbreviated.

  In addition to text, it may contain:

  • **a** • **divineName** • **foreign** • **index** • **name** • **note** • **reference** • **w**

  Attributes of the <abbr> element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **expansion** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **actor:** Actor is used to encode the name of an actor in a **castItem** element, which itself occurs in a **castList** element.

  In addition to text, it may contain:

  • **a** • **abbr** • **foreign** • **index** • **note** • **reference** • **seg** • **w**

  Attributes of the <actor> element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **who** • **xml:lang** • **xml:space**

- **caption:** Caption is used in the **figure** element to record the caption for a map, image or similar items.

  In addition to text, it may contain:

  • **a** • **abbr** • **divineName** • **foreign** • **hi** • **index** • **lb** • **milestone** • **name** • **note** • **q** • **reference** • **seg** • **w**

Attributes of the <caption> element:

• **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **castGroup:** The castGroup element does not allow text content.

  This element contains only:

  • **castItem** • **head**

  Attributes of the **castGroup** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **castItem:** The castItem is a container that groups together information for a particular member of a cast. It does not allow for any text, only the elements listed below.

  This element contains only:

  • **actor** • **role** • **roleDesc**

  Attributes of the **castItem** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **castList:** The castList element appears in the **work** element of an OSIS document to contain one or more castGroup elements. It does not allow any text to appear.

  This element contains only **castGroup** elements.

  Attributes of the **castList** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **catchWord:** The catchWord element is used in **note** and **p** elements that may appear in **note** elements. A catchWord is a portion of text that is repeated, commonly in **bold** typeface to draw the reader's attention to a particular portion of text.

  A **catchWord** element may contain the following elements in addition to text:

  • **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **inscription** • **lb** • **list** • **mentioned** • **milestone** • **name** • **note** • **q** • **reference** • **seg** • **speaker** • **title** • **transChange** • **w**

  Attributes of the <catchWord> element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **osisRef** • **resp** • **script**

• **type** • **subType** • **xml:lang** • **xml:space**

- **cell:** The cell is used in a **row** element of a **table** element to contain content. That content may be text or a mixture of text and the elements listed below. There is an optional **align** attribute that should be used to specify the alignment of cell content. The permitted values are: left, right, center, justify, start and end.

  The **cell** may contain the following elements:

  • **a** • **abbr** • **closer** • **date** • **divineName** • **figure** • **foreign** • **hi** • **index** • **inscription** • **lb** • **lg** • **list** • **milestone** • **name** • **note** • **p** • **q** • **reference** • **salute** • **seg** • **signed** • **speaker** • **speech** • **table** • **title** • **transChange** • **verse** • **w**

  Attributes of the <cell> element:

  • **align** • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **role** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **chapter:** Chapter is used as *syntactic sugar* for the **div** element. It will most often be found in non-biblical texts.

  In addtion to text, the following elements are permitted:

  • **a** • **abbr** • **closer** • **date** • **div** • **divineName** • **figure** • **foreign** • **hi** • **index** • **inscription** • **lb** • **lg** • **list** • **mentioned** • **milestone** • **name** • **note** • **p** • **q** • **reference** • **salute** • **seg** • **signed** • **speaker** • **speech** • **table** • **title** • **transChange** • **verse** • **w**

  Attributes of the **chapter** element:

  • **annotateRef** • **annotateType** • **canonical** • **chapterTitle** • **editions** • **ID** • **n** • **osisID** • **osisRef** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **closer:** The **closer** element is used for the closing portion of a letter, usually containing a final greeting or blessing and a signature.

  In addition to text, the following elements are permitted:

  • **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **lg** • **list** • **mentioned** • **milestone** • **name** • **note** • **p** • **q** • **reference** • **seg** • **signed** • **speaker** • **transChange** • **verse** • **w**

  Attributes of the <closer> element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **contributor:** The **contributor** element appears only in a **work** element. It is used to list a person or organization with a role other than creating the text in question, such as a funder, illustrator, etc. The **role** attribute is used to specify the nature of the role. The content of this element is text only.

  Attributes of the <contributor > element:

• **annotateRef** • **annotateType** • **canonical** • **editions** • **file-as** • **ID** • **n** • **osisID** • **resp** • **role** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **coverage:** The **coverage** element appears only in a **work** element. It is used to specify what is covered by a work, such as a time period, geographic area, etc. The content of this element is text only.

  Attributes of the **coverage** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **creator:** The **creator** element appears only in a **work** element. The person or organization principally responsible for the content of a work is encoded with this element. The **role** attribute is used to specify the nature of that person or organization. The content of this element is text only.

  Attributes of the <creator> element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **role** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **date:** The **date** element is used to record a date in an OSIS document. The **type** attribute is used to indicate the calendar used for the date. Unless otherwise indicated by that attribute, the default calendar is ISO. The other types of calendars permitted by the OSIS schema are: Chinese, Gregorian, Islamic, Jewish and Julian.

  In addition to text, the following elements are permitted:

  • **a** • **index** • **note** • **w**

  Attributes of the <date> element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **event** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **description:** The **description** element is only within a **work** element. It is used to provide a reader friendly description of the work where it appears. The only defined value for the **descriptionType** attribute in the 2.11 version of the OSIS schema is **usfm**. The content of this element is text only.

  Attributes of the **description** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **div:** The **div** element is the basic divider for all OSIS texts.

  The **type** attribute on this element is used to indicate the type of the division. If the **type** attribute is used, its value must either start with "x-" or be one of the following values:

  • **acknowledgement** • **afterword** • **annotant** • **appendix** • **article** • **article** • **back** • **body** •

97

book • **bookGroup** • **chapter** • **colophon** • **commentary** • **concordance** • **coverPage** • **dedication** • **devotional** • **entry** • **front** • **gazetteer** • **glossary** • **imprimatur** • **index** • **introduction** • **majorSection** • **map** • **outline** • **paragraph** • **part** • **preface** • **section** • **subSection** • **titlePage**

In addition to text, the following elements are permitted:

• **a** • **abbr** • **chapter** • **closer** • **date** • **div** • **divineName** • **figure** • **foreign** • **hi** • **index** • **inscription** • **lb** • **lg** • **list** • **mentioned** • **milestone** • **name** • **note** • **p** • **q** • **reference** • **salute** • **seg** • **signed** • **speaker** • **speech** • **table** • **title** • **transChange** • **verse** • **w**

Attributes of the **div** element:

• **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **osisRef** • **resp** • **script** • **type** • **scope** • **subType** • **xml:lang** • **xml:space**

- **divineName:** The **divineName** element is used to mark the name of the Deity only. Other names, such as for angels, demons, idols, etc., should be encoded using the **name** element.

  In addition to text, the following elements are permitted:

  • **a** • **abbr** • **foreign** • **index** • **note** • **reference** • **seg** • **w**

  Attributes of the **divineName** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **figure:** The **figure** element is used to insert maps, images and other materials into an OSIS document. The **src** attribute points to the file to be inserted at the point where the **figure** element occurs ini the text. The **alt** attribute should be used to provide content for readers who may not be able to view graphic images or maps.

  This element does not allow any text content directly, but does permit the following elements:

  • **caption** • **index** • **note**

  Attributes of the **figure** element:

  • **alt** • **annotateRef** • **annotateType** • **canonical** • **catalog** • **editions** • **ID** • **location** • **n** • **osisID** • **osisRef** • **resp** • **rights** • **script** • **type** • **size** • **src** • **subType** • **xml:lang** • **xml:space**

- **foreign:** The **foreign** element is used to mark "foreign" words or phrases in a text. That is words or phrases in a different language from the main text.

  In addition to text, the following elements are permitted:

  • **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **milestone** • **name** • **note** • **reference** • **seg** • **title** • **w**

Attributes of the **foreign** element:

• **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **format:** The **format** element appears only in a **work** element. It is recommended that the format of a work be identified with a standard value for computer media. This element permits only text and no other elements in its content.

  Attributes of the **format** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **head:** The **head** element is used for book, chapter and other headings. While those are sometimes called *titles,* the use of the **head** element allows them to be distinguished from **titles,** which are titles of books or articles.

  In addtion to text, the following elements are permitted:

  • **a** • **abbr** • **date** • **divineName** • **foreign** • **head** • **hi** • **index** • **lb** • **milestone** • **name** • **note** • **q** • **reference** • **seg** • **w**

  Attributes of the **head** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **header:** The **header** element is a container that appears only in the **osisCorpus** or **osisText** elements. When it is used in an **osisCorpus** element, it applies to all the texts in the corpus. When used in an **osisText** element, it applies only to the text of that element.

  No text is permitted in this element, only the following elements:

  • **revisionDesc** • **work** • **workPrefix**

  Attributes of the **header** element:

  • **canonical**

- **hi:** The **hi** element is used to mark text that is typographically distinct from the surrounding text. Text in *italic* for example. If the reason "why" the text is italicized or bolded is apparent, such as a foreign word or phrase, or the name of the Deity, then more specific markup should be used.

  In addtion to text, the following elements are permitted:

  • **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **milestone** • **name** • **note** • **reference** • **seg** • **w**

  Attributes of the **hi** element:

• **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

· **identifier:** The **identifier** element appears only in the **work** element. The value of the **type** attribute for this element must either begin with "x-" or be one of the following systems of identification:
   · Dewey
   · DOI
   · ISBN
   · ISSN
   · LCCN
   · OSIS
   · SICI
   · URI
   · URL
   · URN

This element permits text only.

Attributes of the **identifier** element:

• **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

· **index:** The **index** element has no content and can appear anywhere in the body of an OSIS document. Its purpose is to enable the automatic construction of indexes of OSIS documents. Each **index** element has a required **index** attribute which specifies the index (such as name, subject, etc.) to which this entry belongs, and up to four levels of indexing, the first of which is required if this element is used.

This element does not permit any text or element content. All values associated with it are represented by attribute values.

Attributes of the **index** element:

• **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **index** • **level1** • **level2** • **level3** • **level4** • **n** • **osisID** • **resp** • **script** • **see** • **type** • **subType** • **xml:lang** • **xml:space**

· **inscription:** The **inscription** element is used to mark text that reports a written inscription.

In addition to text, the following elements are permitted:

• **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **milestone** • **name** • **note** • **q** • **reference** • **seg** • **speaker** • **speech** • **transChange** • **w**

Attributes of the **inscription** element:

• **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

· **item:** The **item** element is used within a **list** element to hold the content of each item in the list. An

item may have an optional **label** element that occurs before an content for the purpose of labeling the content of an item.

In addition to text, the following elements are permitted:

• **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lg** • **milestone** • **name** • **note** • **q** • **reference** • **seg** • **title** • **transChange** • **verse** • **w**

Attributes of the **item** element:

• **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **role** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **l:** The **l** element is used to mark separate lines in a **lg** (line group) element. This will be most commonly used in the encoding of poetry.

  The **level** attribute takes a positive integer value that indicates which line in a line group is being encoded. This can be used in processing to insure proper indentation of lines. The **type** attribute, in addition to "x-" takes the following values:

  • **refrain** • **doxology** • **selah** • **attribution**

  In addition to text, the following elements are permitted:

  • **a** • **abbr** • **chapter** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **mentioned** • **milestone** • **name** • **note** • **q** • **reference** • **seg** • **speaker** • **transChange** • **verse** • **w**

  Attributes of the **l** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **level** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **label:** The **label** element is used in an **item** element to provide a label for the content of an item.

  In addition to text, the following elements are permitted:

  • **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **milestone** • **name** • **note** • **q** • **reference** • **seg** • **transChange** • **w**

  Attributes of the **label** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **role** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **language:** The **language** element appears only in a **work** element. There can be multiple **language** elements in a single **work** element.

  If the content of a **language** element identifies the language being used in the OSIS document, the **type** attribute may have an "x-" value or one of the following:

• IANA • IETF • ISO-639-1 • ISO-639-2 • ISO-639-2-B • ISO-639-2-T • LINGUIST • other • SIL

A **language** element also has a **use** attribute which is used to indicate what portion of the text uses the language in question. In addition to a value beginning with "x-", the following values may be used:

• **base** • **didactic** • **interlinear** • **original** • **quotation** • **source** • **target** • **translation**

Attributes of the **language** element:

• **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **use** • **xml:lang** • **xml:space**

- **lb:** The **lb** element is used to indicate a typographical line break in a text. As a milestone type element, it does not have any content. Values for its **type** attribute must begin with an "x-" as there are no predefined types for this element. It can be used to enforce typographical linebreaks for presentation purposes.

  Attributes of the **lb** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **lg:** The **lg** element is generally used to group other **lg** (line group) and **l** (line) elements together.

  The **lg** element does not permit any text as content but only the following elements:

  • **chapter** • **index** • **l** • **lb** • **lg** • **milestone** • **q** • **verse**

  Attributes of the **lg** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **list:** The **list** element is used to represent lists and can include a **head** element for the list.

  In addition to text, the following elements are permitted:

  • chapter (only in milestoneable form) • head: use to provide a heading for **list** • index • item • lb • list: A **list** can hold another **list** • milestone • q • verse

  Attributes of the **list** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **mentioned:** The **mentioned** element is used when a name, for instance, is mentioned but not used as a form of address.

  In addition to text, the following elements are permitted:

• **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **milestone** • **name** • **note** • **q** • **reference** • **seg** • **transChange** • **w**

Attributes of the **mentioned** element:

• **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **milestone:** The **milestone** element is used to mark a location in the text. It does not permit any content.

  The **type** attribute in addition to values starting with "x-", takes the following values:

  • **column** • **cQuote** (continued quote) • **footer** • **halfline** • **header** • **line** • **pb** (page break) • **screen**

  The **marker** attribute takes a value to represent a typographic marker in the original text. This was added to accomodate languages with complex traditions of marking breaks in a text that may not be easily computed.

  Attributes of the **milestone** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **marker** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **milestoneEnd:** This element should not be used in current OSIS documents. It has been replaced by the milestoneable technique which provides better coverage of all known use cases.
- **milestoneStart:** This element should not be used in current OSIS documents. It has been replaced by the milestoneable technique which provides better coverage of all known use cases.
- **name:** The **name** element is used to mark place, personal and other names in an OSIS text. The names of the Deity should always be encoded with **divineName**.

  The **type** attribute, in addition to values starting with "x-" may have one of the following values:

  • **geographic** • **holiday** • **nonhuman** • **person** • **ritual**

  In addition to text, the following elements are permitted:

  • **a** • **abbr** • **foreign** • **index** • **note** • **reference** • **seg** • **w**

  Attributes of the **name** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **regular** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **note:** The **note** element is used for all notes on a text. Liberal use of the **type** attribute will enable easy production of Bible texts with notes appropriate to the reader or situation.

  The **type** attribute, in addition to values beginning with "x-" may have the following values:

• allusion • alternative • background • citation • crossReference • devotional • encoder • exegesis • explanation • liturgical • **speaker** • study • translation • variant

In addition to text, the following elements are permitted:

• **a** • **abbr** • **catchWord** • **date** • **divineName** • **figure** • **foreign** • **hi** • **index** • **inscription** • **lb** • **lg** • **list** • **mentioned** • **milestone** • **name** • **note** • **p** • **q** • **rdg** • **rdgGrp** • **reference** • **seg** • **table** • **title** • **verse** • **w**

Attributes of the **note** element:

• **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **osisRef** • **placement** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **osis:** The **osis** element is the root element of all OSIS texts.

  The **osis** element has no attributes and may contain only one of two other elements: **osisCorpus** or **osisText**.

- **osisCorpus:** The **osisCorpus** element has no attributes and may have a **header**, followed by an optional **titlePage** element, and then one or more **osisText** elements. This element is most useful for collections of separate texts, such as sermons by different ministers encoded as separate OSIS documents or other collections of independent texts.
- **osisText:** The **osisText** element is the main container for a text encoded in OSIS. It is composed of elements only, in the following order: **header** (required), **titlePage** (optional), and one or more **div** elements.

  The **osisIDWork** attribute is required and binds this OSIS text to a particular **work** element in the **header**.

  Attributes of the **osisText** element:

  • **annotateRef** • **canonical** • **ID** • **osisID** • **osisIDWork** • **osisRefWork** • type • **subType** • **xml:lang** • **xml:space**

- **p:** The **p** element is used to mark paragraphs in a text. Since paragraphs are one of the most common features of modern texts, it will probably be the most common element in any translation.

  In addition to text, the following elements are permitted:

  • **a** • **abbr** • **catchWord** • **chapter** (only in milestone form) • **closer** • **date** • **divineName** • **figure** • **foreign** • **hi** • **index** • **inscription** • **lb** • **lg** • **list** • **mentioned** • **milestone** • **name** • **note** • **q** • **rdg** • **rdgGrp** • **reference** • **salute** • **seg** • **signed** • **speaker** • **speech** • **table** • **title** • **transChange** • **verse** • **w**

  Attributes of the **p** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **publisher:** The **publisher** element occurs only in a **work** element.

  If a particular text has been published several times, there should be separate **work** elements for each separate publication. The publication that the OSIS text is based upon should be linked by using the **osisIDWork** attribute from the appropriate **work** element as the value of the **osisIDWork** attribute on the **osisText** element for the text.

  Attributes of the **publisher** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **q:** The **q** element is used to mark quotations in a text.

  The **who** attribute should be liberally used to support searching for quotations by a particular person or rendering the words of Jesus in red, for example. Due to the complexity of marking quotations, the **marker** attribute allows the recording of a particular character to mark the quotation. The **level** attribute should be used in cases where there are multiple quotations embedded in each other.

  In addition to text, the following elements are permitted:

  • **a** • **abbr** • **closer** • **date** • **divineName** • **foreign** • **hi** • **index** • **inscription** • **lb** • **lg** • **list** • **mentioned** • **milestone** • **name** • **note** • **p** • **q** • **reference** • **salute** • **seg** • **signed** • **speaker** • **transChange** • **verse** • **w**

  Attributes of the **q** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **level** • **marker** • **n** • **osisID** • **osisRef** • **resp** • **script** • **type** • **subType** • **who** • **xml:lang** • **xml:space**

- **rdg:** The **rdg** element is used to record a variant reading of a text. Most often seen where a note says: "Hebrew/Greek reads:".

  The **type** attribute, in addition to values starting with "x-", takes the following values:

  • **alternative** • **variant**

  In addition to text, the following elements are permitted:

  • **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **lg** • **list** • **milestone** • **name** • **note** • **p** • **q** • **reference** • **seg** • **transChange** • **w**

  Attributes of the **rdg** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **witness** • **xml:lang** • **xml:space**

- **rdgGrp** The **rdgGrp** element is used to group together **rdg** elements. The only content allowed is one or more **rdg** elements.

Attributes of the **rdgGroup** element:

• **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **reference:** The **reference** element is used to mark references in one text to another text. The **type** attribute takes a value that conforms to an OSIS reference.

  • **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **milestone** • **name** • **note** • **seg** • **title** • **w**

  Attributes of the **reference** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **osisRef** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **refSystem:** The **refSystem** element occurs only in a **work** element. It has text only content and is used to specify for a reader the reference system used by a particular text.

  Attributes of the **refSystem** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **relation:** The **relation** element occurs only in a **work** element. It has only text only content and is used to specify for a reader the relationship of one work to another.

  Attributes of the **relation** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **revisionDesc:** The **revisionDesc** element is used only in a **header** element. It is used to record changes to a text as it is encoded. Liberal use of this element is encouraged.

  Each **revisionDesc** element consists of one **date** element, followed by one or more **p** elements.

  Attributes of the **revisionDesc** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **rights:** The **rights** element is used only in a **work** element. It is used to specify for a reader the intellectual property rights asserted against a particular text. Such as a copyright by a publisher or translator.

  Attributes of the **rights** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

106

- **role:** Role is used in a **castItem** element to identify the role of a particular actor.

  In addtion to text, it may contain:

  • **a** • **abbr** • **foreign** • **index** • **note** • **reference** • **seg** • **w**

  Attributes of the **role** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **roleDesc:** The roleDesc element is used to provide a description of a role in a **castItem** element.

  In addition to text, it may contain:

  • **a** • **abbr** • **foreign** • **index** • **note** • **reference** • **seg** • **w**

  Attributes of the **roleDesc** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **row:** The **row** element occurs only in **table** elements and is used to contain **cell** elements.

  Attributes of the **row** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **role** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **salute:** The **salute** element is used to mark a saluation or opening comments. It is most generally associated with letters and similar materials.

  In addition to text, the following elements are permitted:

  • **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **lg** • **list** • **mentioned** • **milestone** • **name** • **note** • **p** • **q** • **reference** • **salute** • **seg** • **speaker** • **transChange** • **verse** • **w**

  Attributes of the **salute** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **scope:** The **scope** element is used only in a **work** element. The general area covered by a text is indicated to a reader with this element. It has only text content.

  Attributes of the **scope** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

107

- **seg:** The **seg** element should be used for very small divisions, such as within **word** elements. The varieties of morphological divisions in languages made it impossible for the OSIS editors to anticipate every division that would be useful to encoders.

  The **type** attribute can have a value that starts with "x-" or one of the following types:

  • **alluded** • **keyword** • **otPassage** • **verseNumber**

  In addition to text, the following elements are permitted:

  • **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **milestone** • **name** • **note** • **q** • **reference** • **seg** • **transChange** • **w**

  Attributes of the **seg** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **signed:** The **signed** element is used to mark the signer of a letter within a **closer** element.

  In addition to text, the following elements are permitted:

  • **a** • **abbr** • **divineName** • **foreign** • **hi** • **index** • **lb** • **milestone** • **name** • **note** • **reference** • **seg** • **w**

  Attributes of the **signed** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **source:** The **source** element appears only in a **work** element. It is used to indicate the source for a text that is being encoded. It consists of text only.

  Attributes of the **source** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **speaker:** The **speaker** element is used to mark the speaker in a text. It will be used when the speaker is named in the text. If the speaker is not named in the text, the **who** attribute on the **q** element should be used.

  In addition to text, the following elements are permitted:

  • **a** • **abbr** • **divineName** • **foreign** • **index** • **name** • **note** • **reference** • **w**

  Attributes of the **speaker** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **who** • **xml:lang** • **xml:space**

- **speech:** The **speech** element is used to mark speeches in a text.

  The **marker** attribute can be used to indicate how a speech is delimited in a particular style of text.

  In addition to text, the following elements are permitted:

  • **a** • **abbr** • **chapter** • **closer** • **date** • **divineName** • **foreign** • **hi** • **index** • **inscription** • **lb** • **lg** • **list** • **mentioned** • **milestone** • **name** • **note** • **p** • **q** • **reference** • **salute** • **seg** • **signed** • **speaker** • **table** • **title** • **transChange** • **verse** • **w**

  Attributes of the **speech** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **marker** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **subject:** The **subject** element occurs only in a **work** element. It consists only of text drawn from a specified system of subject classification.

  The **type** attribute, in addition to values starting with "x-" can have one of the following values:

  • **ATLA**: American Theological Library Association • **BILDI**: Biblewissenschaftliche Literaturdokumentation Innsbruck • **DBC**: Dutch Basic Classification • **DDC**: Dewey Decimal Classification • **EUT**: Estonian Universal Thesaurus • **FGT**: Finnish General Thesaurus • **LCC**: Library of Congress Subject Heading • **MeSH**: Medical Subject Headings • **NLSH**: National Library Subject Headings (National Library of Poland) • **RSWK**: Regeln für den Schlagwortkatalog • **SEARS**: Sears List of Subject Headings • **SOG**: Soggettario • **SWD_RSWK**: Swiss National Library • **UDC**: Universal Decimal Classification • **VAT**: Vatican Library

  Attributes of the **subject** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **table:** The **table** element contains an optional **head** element and one or more **row** elements.

  As an aid to processing and display, the optional attributes **cols** and **rows** are available and both take positive integer values.

  Attributes of the **table** element:

  • **annotateRef** • **annotateType** • **canonical** • **cols** • **editions** • **ID** • **n** • **osisID** • **resp** • **rows** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **teiHeader:** The **teiHeader** element occurs only in a **header** element. It is used to contain a TEI header for the text that is being encoded.

  The **teiHeader** element has no attributes that allow user input.

- **title:** The **title** element is used to record a title both in a **work** element and elsewhere in an OSIS text.

Due to the complexities of presentation of titles, there are several attributes that can assist in the processing of **title** elements.

- The **level** attribute takes an integer to indicate the level of the title.
- The **short** attribute takes a short form of the title.
- The **placement** attribute takes a value starting with "x-" or one of the following values:

  • **leftHead** • **centerHead** • **rightHead** • **insideHead** • **outsideHead** • **leftFoot** • **centerFoot** • **rightFoot** • **insideFoot** • **outsideFoot**

- The **type** attribute takes one of the following values:

  • **acrostic** • **chapter** • **continued** • **main** • **parallel** • **runningHead** • **scope** • **sub**

In addition to text, the following elements are permitted:

• **a** • **abbr** • **date** • **divineName** • **figure** • **foreign** • **hi** • **index** • **inscription** • **lb** • **lg** • **milestone** • **name** • **note** • **q** • **reference** • **seg** • **title** **verse** • **w**

Attributes of the **title** element:

• **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **level** • **n** • **osisID** • **placement** • **resp** • **script** • **short** • **type** • **subType** • **xml:lang** • **xml:space**

- **titlePage:** The **titlePage** element is used to specify a particular title page for an OSIS document. Normally that should occur through the use of XSLT transformation software based on the content of the text. In exceptional circumstances, however, the **titlePage** element can be used to encode a static representation of a title page.

  Only the following elements are permitted (in any combination):

  • **title** • **contributor** • **creator** • **subject** • **date** • **description** • **publisher** • **type** • **format** • **identifier** • **source** • **language** • **relation** • **coverage** • **figure** • **milestone** • **p**

  Attributes of the **titlePage** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **transChange:** The **transChange** element is used to mark text that is not present in the original language. Such as words added to smooth out a translation where the original language has a different syntax structure from the original.

  The **type** attribute accepts values starting with "x-" or one of the following values:

  • **added** • **amplified** • **changed** • **deleted** • **implied** • **moved** • **tenseChange**

  In addition to text, the following elements are permitted:

  • **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **lb** • **milestone** • **name** • **note** • **q** •

reference • **seg** • **w**

Attributes of the **transChange** element:

• **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **type:** The **type** element occurs only in a **work** element. It is used to indicate to the reader the type of the text.

  Attributes of the **type** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

  Attributes of the **type** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **verse:** The **verse** element should almost always be used in its milestoneable form. While some older texts can be represented using the traditional start/end tags for **verse** elements, in the long run that diminishes the usefulness of the text.

  In addition to text, the following elements are permitted:

  • **a** • **abbr** • **date** • **divineName** • **foreign** • **hi** • **index** • **inscription** • **lb** • **list** • **mentioned** • **milestone** • **name** • **note** • **q** • **reference** • **seg** • **speaker** • **title** • **transChange** • **w**

  Attributes of the **verse** element:

  • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

- **w:** The **w** element is used to encode particular words in a text.

  Because of the important of words in biblical study, the following optional attributes are available on the **w** element:

  - **gloss**: Use for short explanation of a word.
  - **lemma**: Use for the base form of a word.
  - **morph**: Use to record a word's morphology.
  - **POS**: Use to record part of speech analysis.
  - **src**: Use for the source of a word.
  - **xlit**: Use for a transliterated form of the word.

  In addition to text, the following elements are permitted:

  • **a** • **index** • **note** • **seg**

Attributes of the **w** element:

> • **annotateRef** • **annotateType** • **canonical** • **editions** • **gloss** • **ID** • **lemma** • **morph** • **n** • **osisID** • **POS** • **resp** • • **script** • **type** • **subType** • **xml:lang** • **xml:space** • **xlit**

- **work:** The **work** element occurs only in a **header** element. It provides all the basic identification and other information for a text that is being encoded in OSIS. It should be as complete as possible.

  Only the following elements are permitted:

  > • **title** • **contributor** • **creator** • **subject** • **date** • **description** • **publisher** • **type** • **format** • **identifier** • **source** • **language** • **relation** • **coverage** • **rights** • **scope** • **castList** • **teiHeader** • **refSystem**

  Attributes of the **work** element:

  > • **annotateRef** • **annotateType** • **canonical** • **editions** • **ID** • **n** • **osisID** • **osisWork** • **resp** • **script** • **type** • **subType** • **xml:lang** • **xml:space**

Appendix B Alphabetical list of Attributes and normative values

Appendix B.1 Global attributes

Global attributes are attributes that are available on every element. Some elements have specific attributes that are listed separately by each element below. The following global attributes are in addition to the attributes xml:base, xml:lang, and xml:space, which are defined by the XML standard itself.

Attribute name	DataType	Usage	Description
annotateRef	xs:string	optional	The value of an **annotateRef** attribute is an **osisRef** that points to a portion of a work that is being annotated by the element where the **annotateRef** appears.
annotateType	osisAnnotation	optional	The **annotateType** attribute specifies for the **annotateRef** on the element where it appears, the type of annotation that is being made. It is an OSIS error for the **annotateType** attribute to have a value when a value has not been specified for the **annotateRef** attribute.
annotateWork	xs:string	optional	The **annotateWork** attribute has been depracated. (That means don't use it.) Use **annotateRef** and **annotateType** instead.
canonical	true | false	optional	The canonical attribute identifies the element bearing it as containing actual text of the work being encoded, as opposed to annotations, commentary, inserted headings, header metadata, notes, and other (non-canonical) information. Its value inherits in the same way as xml:lang. That is, the value applies to all descendant elements except where overridden.
editions	xs:nmtokens	optional	The **editions** attribute is used to specify edition specific content, such as material omitted in Protestant versions of the Bible, for instance.
ID	xs:ID	optional	May be added to any element, generally to make that element accessible as a link target for generic hypertext browsers. Duplicate ID's are not permitted in the same document.
n	xs:string	optional	This attribute is identical to the TEI n attribute, and may be used to provide a name or number to identify the particular element instance. However, it should not be used to encode a value for which the **osisID**, **osisRef**, or other attribute is

applicable.

osisID	osisIDType	optional	The osisID attribute identifies the element bearing it as a container for actual canonically-referenceable text, and provides the applicable osisID: osisID="Matt.1.1". It must not be used on elements that merely refer to, or discuss, a canonically-referenceable text. For those cases, use the annotateWork and osisRef attributes, instead. See the section on reference systems for details on the form required for this attribute's value.
osisRef	osisRefType	optional	The osisRef attribute points to any reasource that can be identified by an osisID. This will most often be used to point to other biblical texts that have been encoded using OSIS, but it can point to any text encoded using OSIS.
resp	xs:string	optional	This attribute, whose name is short for "responsible party", may be coded on any element to identify the party primarily responsible for that element and its content. For example, it might identify a member of a translation team; or on a note, it might identify the author of the note. Each distinct responsible party must be identified by the same value of this attribute wherever they are identified at all (that is, it is not permitted to use their initials sometimes, their last names other times, etc.).
script	osisScripts	optional	This attribute provides a slight extension beyond the capabilities of the **xml:lang** attribute. For many languages, it is enough to specify the language itself, and the country where it is spoken (say, Canadian vs. French dialects of the French language). However, there are cases where a given language community may use multiple writing systems: either different character sets and character usage rules; different spelling or other conventions; etc. In such cases, the particular script system used for writing the current work (or element within a work) must be specified via this attribute. This attribute inherits in precisely the same manner as xml:lang.
subType	xs:string	optional	In the rare event that the **type** attribute does not provide a fine-enough grained distinction of element types, the sub-type attribute may be used to make such distinctions. There are generally no restrictions on the values permitted for subType, except that the encoder should be consistent, and should document the meaning of any values used.
type	xs:string (several element types	optional	The type attribute allows encoders to identify more precise distinctions within the broad applicability of any given element. For example, the **div** (division) element has many

114

	restrict the values, as listed below).		subtypes, such as bookGroup, concordance, dedication, glossary, etc. Many other element types also have pre-defined values provided for the type attribute. Some but not all of those element types also permit users to add their own values, so long as they begin with "x-". When a predefined type is applicable, it must be used instead of creating a new type.
xml:lang	xs:language	optional	This attribute is defined by the XML standard and identifies the primary natural language of the content of an element. The value of this attribute is inherited; that is, any contained elements are presumed to be in the same language, unless they override it by specifying their own explicit **xml:lang** attribute value. The form of the **xml:lang** attribute is constrained by Internet specifications, particularly IETF RFC 1766, **Tags for the Identification of Languages.** Such tags consist of a 2-letter language code from ISO 639 (see http://www.oasis-open.org/cover/iso639a.html), optionally followed by a two letter country code from ISO 3166 (see http://www.oasis-open.org/cover/country3166.html). For example, "de" or "en-GB" Alternatively, they may be codes from the IANA registry at http://www.iana.org/assignments/language-tags. Remaining languages should use SIL Ethnologue codes (see http://www.ethnologue.com/codes/).
xml:space	xs:space	optional	This attribute is defined by the XML standard and indicates to an XML processor whether whitespace between elements should be preserved or not. This can be important when encoding poetry and other material where precise layout needs to be specified.

Appendix B.2 Normative values for the type attribute, by element

The heading for each basic element type below, indicates whether the list of values provided is extensible (by adding names beginning with "x-", or non-extensible).

OSIS is an evolving encoding standard designed to meet the needs of a wide variety of user communities. Should you discover attribute values that may be of general interest, please forward them to the editors for possible inclusion in a future OSIS release.

Appendix B.2.1 annotateType

The **annotateType** attribute, which is available on all content elements, is to be used along with the **annotateRef** attribute, to indicate the type of annotation is being made to another work or portion of another work. This sort of reference can point to anything that can be pointed to using an **osisRef** so the annotation about a word, phrase, verse, chapter or larger portion of text.

These values characterize the annotation, not the work that is being annotated. They will be particularly helpful in systems where annotations of a particular type, **rebuttal** for example, are being sought for a particular work.

If the type of annotation you are making does not appear in this list, use the OSIS attribute extension, that is, "x-" in front of your attribute value.

- **commentary** A comment or fuller commentary on the reference given by the **annotateRef** attribute.
- **exposition** A development of the meaning of the reference given by the **annotateRef** attribute..
- **meditation** A meditation on the reference given by the **annotateRef** attribute..
- **outline** An outline of the reference given by the **annotateRef** attribute..
- **rebuttal** A rebuttal of one or more points in the reference given by the **annotateRef** attribute..
- **sermon** A sermon based on the reference given by the **annotateRef** attribute.
- **studyGuide** A study guide on the reference given by the **annotateRef** attribute.
- **translation** A translation of the reference given by the **annotateRef** attribute.

Appendix B.2.2 date

What is considered to be the standard calendar varies by historical time period as well as culture. The **type** attribute on the **date** element specifies what calendar system was used for the content of the **date** element.

For cases where the required calendar is not one of the following values, please use the OSIS attribute extension mechanism, "x-" in front of the name of your calendar.

- **Chinese** Information on the Chinese calendar can be found at: http://webexhibits.org/calendars/calendar-chinese.html
- **Gregorian** Standard calendar in use in the US and Europe. For further information see: http://www.geocities.com/calendopaedia/gregory.htm
- **Islamic** Standard calendar in Muslim countries and religious communities. For further information see: http://webexhibits.org/calendars/calendar-islamic.html
- **ISO** This is not a separate calendar from the Gregorian but is a specific notation for recording dates. For further information see: http://www.cl.cam.ac.uk/~mgk25/iso-time.html
- **Jewish** Official calendar of Israel and for religious purposes. For further information see: http://webexhibits.org/calendars/calendar-jewish.html
- **Julian** Largely historical but note that the Julian calendar continues in use by the Russian Orthodox Church. For further information see: http://www.geocities.com/calendopaedia/julian.htm

Appendix B.2.3 div

The **type** attribute for **div** mainly identifies larger sections that occur in print volumes, especially Bibles. This list was determined in part by examining a large selection of print Bibles, and covers most things that seem to be common. However, the list may be extended if necessary by adding names beginning "x-".

- **acknowledgement** Standard acknowledgement page.
- **afterword** The afterword in a text.
- **annotant** Signals an annotation of another text.
- **appendix** The appendix, in Bibles usually where maps, tables and similar material is found.
- **article** An short work, usually by a separate author from the main work, on a particular theme or

topic.

- **back** Similar to appendix. Legacy of generic encoding systems that divide texts into front, body and back material.
- **body** The main portion of a text.
- **book** Bibles generally refer to the **Book** of Genesis. While at variance with standard publishing terminology, there is no good reason to not continue this tradition.
- **bookGroup** Should be used in Bibles to contain groups of **book**s, such as gathering all the books of the Torah or similar groups.
- **chapter** Standard chapter as is found in a textbook.
- **colophon** Most often found in hand written texts to identify the author, place of composition but does occur in some printed works.
- **commentary** Useful for separating divisions in a text that have blocks of Bible text from blocks of text that are commentaries on the Bible text.
- **concordance** Should be used for short concordances that are bound with Bibles and other works. Full concordances are texts in their own right.
- **coverPage** A page in printed books before the actual title page.
- **dedication** A page, usually in gift Bibles, that record the gift of the Bible to a particular individual.
- **devotional** Should be used to indicate that the division contains a short worship service.
- **entry** Best used in the sense of an entry in a dictionary.
- **front** Typically contains material like a presentation page ('Presented to:'), abbreviations, etc.
- **gazetteer** A geographic dictionary.
- **glossary** A listing of terms with short definitions. Distinguished from a dictionary by the brevity of the entries.
- **imprimatur** Latin for 'let it be printed.' Imprimatur is used to indicate a review, traditionally by an ecclesiastical authority, that certifies the work in question conforms to a particular orthodoxy.
- **index** A listing of words or concepts in a work, usually with page or other references to locations in the work where those words or concepts may be found.
- **introduction** An overview of material about to be covered in the main part of a work.
- **majorSection** A rather loose term used to distinguish large sections of a work from smaller ones. The New and Old Testaments could be considered **majorSection**s of the Bible as a whole.
- **map** A portion of a work that contains maps.
- **outline** A portion of a work that contains an outline, usually for material immediately following this division.
- **paragraph** Should be used only where the paragraphs in question do not correspond to normal use of the **p** element of OSIS.
- **part** A division of a work that is difficult to distinguish from **majorSection**, **section** or **minorSection** but it does occur in practice and users are cautioned to simply be consistent within a work with their usage.
- **preface** Prefaces are commonly the preliminary remarks by an author. May also be used for the opening of a traditional eucharistic liturgy.
- **section** A division of a work, typically used with **majorSection** and **minorSection**
- **subSection** A division of a work, typically used with **majorSection** and **section**.
- **summary** A condensed version of what has just been stated, usually at some length, particularly in biblical or theological commentaries.
- **titlePage** A page in a publication that gives title, author and other related information. Generally speaking the **titlePage** element should be used instead in OSIS texts.

Appendix B.2.4 Identifier

The **type** attribute on the **identifier** element is used to identify the identifier system from which the identifier was obtained. Note that the values for the **type** attribute must be entered exactly as shown, all others must use the "x-" extension mechanism.

If the user uses an identifier system that is not covered by these values, please use the OSIS attribute extension mechanism, "x-" in front of the name of your value for this attribute.

- **Dewey** Dewey Decimal System
- **DOI** Digital Object Identifier
- **ISBN** International Standard Book Number
- **ISSN** International Standard Serial Number
- **LCCN** Library of Congress Control Number
- **OSIS** Open Scriptural Information Standard
- **SICI** Serial Item and Contribution Identifier
- **URI** Uniform Resource Identifier
- **URL** Uniform Resource Locator
- **URN** Uniform Resource Name

Appendix B.2.5 language

The **language** element has two attributes with enumerated values. There is the **type** attribute as well as **use** attribute. The values for both are enumerated below.

Appendix B.2.5.1 type attribute on <language>

The **type** attribute on the **language** element is used to identify the authority for the value found in the element content. This provides the user of the file a reference that will assist them in determining the meaning of the content of the element. Different authorities have different names for the same languages and knowing which authority issued a particular name can be quite helpful.

If the user uses an authority that is not covered by these values, please use the OSIS attribute extension mechanism, "x-" in front of the name of your value for this attribute.

- **IANA** Internet Assigned Numbers Authority: Language codes from the IANA registry: http://www.iana.org/assignments/language-tags.
- **IETF** Internet Engineering Task Force: http://www.ietf.org/rfc/rfc1766.txt
- **ISO-639-1** International Organization for Standardization: http://www.oasis-open.org/cover/iso639.html
- **ISO-639-2** International Organization for Standardization: http://lcweb.loc.gov/standards/iso639-2/langhome.html (from the Library of Congress website)
- **ISO-639-2-B** International Organization for Standardization: (same list as ISO-639-2 but in bibliographic order.)
- **ISO-639-2-T** International Organization for Standardization: (same list as ISO-639-2 but in terminology code order.)
- **LINGUIST** Linguist List: A resource maintained by linguists, http://www.linguistlist.org/.
- **other** Does not use the "x-" extension, simply put "other."

- **SIL** SIL International (formerly, Summer Institute for Linguistics): http://www.ethnologue.com/codes/ (the definitive site for language codes).

Appendix B.2.5.2 use attribute on <language>

The **use** attribute on the **language** element is used to identify how the language identified in the **language** element is used in the text.

The values that are duplicates, that is: **original** and **source** or **target** and **translation** are not errors. Communities have grown up around using one term or the other and since they are equivalent, there appeared to be no reason to arbitrarily pick one over the other. Users should be consistent in their usage within a document and if possible, within collections of documents.

If the user needs a use of language that is not covered by these values, please use the OSIS attribute extension mechanism, "x-" in front of the name of your value for this attribute.

- **base** Language is used as the primary language of the text.
- **didactic** Language is used for instruction, such as in a grammar, in the text.
- **interlinear** As the name implies, one of two or possibly more languages used to represent a text, such as English and Latin. The Loeb translation series being a good example.
- **original** Use where the language is used for the text that is being translated. Duplicate of "source" value below.
- **quotation** Language used for quotations in the text.
- **source** Use where the language is used for the text that is being translated. Duplicate of "original" above.
- **target** Use for the language into which a translation is being made. Duplicate of "translation" below.
- **translation** Use for the language into which a translation is being made. Duplicate of "target" above.

Appendix B.2.6 milestone

The **type** attribute on the **milestone** element is used to identify what type of point is being indicated. Note that the values for the **type** attribute must be entered exactly as shown, all others must use the "x-" extension mechanism.

If the user needs to indicate point in the text that is not covered by these values, please use the OSIS attribute extension mechanism, "x-" in front of the name of your value for this attribute.

- **column** Marks the end of a column where there is a multi-column display.
- **footer** Marks the footer region of a page.
- **halfLine** Used to mark half-line units if not otherwise encoded.
- **header** Marks the header region of a page.
- **line** Marks line breaks, particularly important in recording appearance of an original text, such as a manuscript.
- **pb** Marks a page break in a text.
- **screen** Marks a preferred place for breaks in an on-screen rendering of the text.

Appendix B.2.7 name

The **type** attribute on the **name** element is used to identify the type of name that is being encoded. This is important for searching software to be able to distinguish geographic names from personal names or to create lists of particular types of names. Note that the values for the **type** attribute must be entered exactly as shown, all others must use the "x-" extension mechanism.

If the user needs to record a type of name in the text that is not covered by these values, please use the OSIS attribute extension mechanism, "x-" in front of the name of your value for this attribute.

- **geographic** Name of a place or location.
- **holiday** Name of a holiday or festival.
- **nonhuman** Name of any nonhuman other than the Deity. For the latter, see divineName.
- **person** Name of any person.
- **ritual** Name of a ritual.

Appendix B.2.8 notes

The **type** attribute on the **note** element is used to identify the type of note that appears in the text. This is important for searching software to be able to find notes of a particular type, or to suppress notes that are not of interest for a particular reader or purpose. Note that the values for the **type** attribute must be entered exactly as shown, all others must use the "x-" extension mechanism.

If the user needs to record a type of note in the text that is not covered by these values, please use the OSIS attribute extension mechanism, "x-" in front of the name of your value for this attribute.

- **allusion** Use for a note that records an allusion in the text. Ex: "Call me Ishmael" in a modern novel is an allusion to Moby Dick by Herman Melville. The text does not cite Melville but the encoder notes the allusion.
- **alternative** Use for a note that provides an alternative to the main text, usually with a translation. Differs from variant, see below, in that variant is of the original text, not the translation.
- **background** Use for a note that provides background information, usually about unfamiliar practices, terms or measures.
- **citation** Use for a note that contains a formal citation to another work, modern footnote is a good example.
- **crossReference** Use for a note that provides a cross reference to material relevant to the issue under discussion in the text.
- **devotional** Use for a note that details devotional aspects of text or issues in the text under discussion.
- **exegesis** Use for a note that makes the case for a particular interpretation of a text.
- **explanation** Use for a note that provides a brief explanation of a term or phrase.
- **study** Use for notes that are of particular interest to students (formal or self-directed).
- **translation** Use for notes that comment on issues in a translation or particular translation decisions.
- **variant** Use for notes that provide alternatives to the original text that underlies a particular translation. An example of poor variant practice: 'Some Arabic manuscripts read...' If you provide a variant, please provide a formal citation. To do otherwise, simply annoys the reader.

Appendix B.2.9 subject

The following are the valid values for the **type** attribute on the **subject** element. Note that what is entered is in **bold** and the following material is just for the convenience of the reader. Note that an XML parser will expect the values to be entered exactly as you see them in this list. Case, that is upper or lower, matters to an XML parser. An attribute with the value **ATLA** is VALID, but one with the value **atla** is INVALID. You have been warned.

- **ATLA** American Theological Libraries Association
- **BILDI** Bibelwissenschaftliche Literaturdokumentation Innsbruck
- **DBC** Dutch Basic Classification
- **DDC** Dewey Decimal Classification
- **EUT** Estonian Universal Thesaurus
- **FGT** Finnish General Thesaurus
- **LCC** Library of Congress Classification
- **LCSH** Library of Congress Subject Heading
- **MeSH** Medical Subject Headings
- **NLSH** National Library Subject Headings (National Library of Poland)
- **RSWK** Regeln für den Schlagwortkatalog
- **SEARS** Sears List of Subject Headings
- **SOG** Soggettario
- **SWD_RSWK** Swiss National Library
- **UDC** Universal Decimal Classification
- **VAT** Vatican Library

Appendix B.2.10 titles

The **type** attribute on the **title** element is used to allow special rendering of particular titles, as well as searching for particular types of titles in the text.identify the type of note that appears in the text. Note that the values for the **type** attribute must be entered exactly as shown, all others must use the "x-" extension mechanism.

If the user needs to record a type of title in the text that is not covered by these values, please use the OSIS attribute extension mechanism, "x-" in front of the name of your value for this attribute.

- **acrostic** Use for titles where alignment of first or final letters of words in the title are meaningful.
- **continued** Use for titles that are continuations of some other part of the title.
- **main** Use for the main title of a work.
- **parallel** Use where titles are given in alternative languages.
- **psalm** Use in the Psalms where what are considered "titles" in the English text are actually numbered as verses in the Hebrew text.
- **sub** Use for subparts of a title.

Appendix B.2.11 transChange

The **type** values on the **transChange** element is used to indicate a departure from a literal rendering of the source text. This happens most often when words are added to a translation to make the meaning of the text clearer or when the grammatical structures of the translation language do not offer same tenses for example,

as the source language.

If the user encounters a change in translation that is not covered by these values, use the OSIS attribute extension mechanism, "x-" in front of the name of your value for this attribute.

- **added** Words added.
- **amplified** More than addition of words to smooth out a translation.
- **changed** Words are changed in the translation, such as modern spellings.
- **deleted** Words that appear in the original but not in the translation.
- **moved** Words that are moved to better represent the meaning of the text being translated from their original order.
- **tenseChange** Indicates a change of the tense from the original to one that occurs in the translation language.

Appendix C Normative Abbreviations for canonical and deutero-canonical books

These names are taken from the SBL Manual of Style, which also provides normative abbreviations for works of classical literature, manuscripts, journals, and other information objects of interest to Biblical studies.

Appendix C.1

Hebrew Bible/Old Testament

- **Gen** Genesis
- **Exod** Exodus
- **Lev** Leviticus
- **Num** Numbers
- **Deut** Deuteronomy
- **Josh** Joshua
- **Judg** Judges
- **Ruth** Ruth
- **1Sam** 1 Samuel
- **2Sam** 2 Samuel
- **1Kgs** 1 Kings
- **2Kgs** 2 Kings
- **1Chr** 1 Chronicles
- **2Chr** 2 Chronicles
- **Ezra** Ezra
- **Neh** Nehemiah
- **Esth** Esther
- **Job** Job
- **Ps** Psalms
- **Prov** Proverbs
- **Eccl** Ecclesiastes
- **Song** Song of Solomon
- **Isa** Isaiah
- **Jer** Jeremiah
- **Lam** Lamentations
- **Ezek** Ezekiel
- **Dan** Daniel
- **Hos** Hosea
- **Joel** Joel
- **Amos** Amos
- **Obad** Obadiah
- **Jonah** Jonah
- **Mic** Micah
- **Nah** Nahum
- **Hab** Habakkuk

- **Zeph** Zephaniah
- **Hag** Haggai
- **Zech** Zechariah
- **Mal** Malachi

New Testament

- **Matt** Matthew
- **Mark** Mark
- **Luke** Luke
- **John** John
- **Acts** Acts
- **Rom** Romans
- **1Cor** 1 Corinthians
- **2Cor** 2 Corinthians
- **Gal** Galatians
- **Eph** Ephesians
- **Phil** Philippians
- **Col** Colossians
- **1Thess** 1 Thessalonians
- **2Thess** 2 Thessalonians
- **1Tim** 1 Timothy
- **2Tim** 2 Timothy
- **Titus** Titus
- **Phlm** Philemon
- **Heb** Hebrews
- **Jas** James
- **1Pet** 1 Peter
- **2Pet** 2 Peter
- **1John** 1 John
- **2John** 2 John
- **3John** 3 John
- **Jude** Jude
- **Rev** Revelation

Apocrypha and Septuagint

- **Bar** Baruch
- **AddDan** Additions to Daniel
- **PrAzar** Prayer of Azariah
- **Bel** Bel and the Dragon
- **SgThree** Song of the Three Young Men
- **Sus** Susanna
- **1Esd** 1 Esdras
- **2Esd** 2 Esdras
- **AddEsth** Additions to Esther
- **EpJer** Epistle of Jeremiah
- **Jdt** Judith
- **1Macc** 1 Maccabees

- **2Macc** 2 Maccabees
- **3Macc** 3 Maccabees
- **4Macc** 4 Maccabees
- **PrMan** Prayer of Manasseh
- **Sir** Sirach/Ecclesiasticus
- **Tob** Tobit
- **Wis** Wisdom of Solomon

These abbreviations are as defined in the *SBL Handbook of Style* published by the Society of Biblical Literature, except that spaces have been removed from the abbreviations for some Apocryphal and Septuagint books.

Note that because XML prohibits digits as the first character of IDs and other XML names, these abbreviations cannot be used directly as XML IDs.

Appendix D Standard OSIS Codes for Bible Editions

All Bible Edition codes must have the language code for the target language in question, then a colon, then the abbreviation shown here.

Appendix D.1 Ancient language editions

- **Steph** Stephanus GNT, 1551
- **Vul** Latin Vulgate, 1405
- **Erasmus** Latin translation by Desiderius Erasmus Roterodamus, 1516
- **MT** Masoretic text (various, ~900-1100)
- **BHS** Biblia Hebraica Stuttgartensia
- **NA** Nestle-Aland Greek New Testament (may suffix edition number, such as "NA27")
- **LXX** Greek Septuagint

Appendix D.1.1 English Editions (prefix "en:")

- **AAT** The Complete Bible: An American Translation, by Edgar Goodspeed and J. M. Powis Smith, 1939.
- **ABT** The Afro Bible Translation
- **ATB** The Alternate Translation Bible
- **ASV** American Standard Version
- **AB** The Amplified Bible
- **ALT** Analytical-Literal Translation
- **ASL** American Sign Language Translation
- **AV** Authorized Version (same as KJV)
- **Bar** The New Testament: A New Translation, by William Barclay
- **BB** The Biker Bible
- **BWE** Bible in WorldWide English
- **CCB** Christian Community Bible
- **COM** The Common Edition: New Testament
- **COV** Covenant Edition New Testament
- **CJB** Complete Jewish Bible
- **CONC** Concordant Version
- **CEV** Contemporary English Version
- **CPV** Cotton Patch Version, tr. Clarence Jordan
- **Dar** Darby
- **DR** Douay-Rheims
- **DRP** David Robert Palmer's translations of the gospels
- **EMTV** English Majority Text Version
- **ENT** Extreme New Testament
- **ERV** Easy-to-Read Version
- **ESV** English Standard Version
- **FF** Ferrar Fenton Bible
- **GLW** God's Living Word
- **GNC** God's New Covenant: A New Testament Translation, by Heinz W. Cassirer

- **GW** God's Word
- **GNB** Good News Bible (TEV)
- **HCSB** Holman Christian Standard Bible
- **ICB** International Children's Bible (children's version of the NCV)
- **ISB** International Standard Bible (formerly titled The Simple English Bible)
- **ISV** The International Standard Version
- **JBP** New Testament in Modern English, by J.B. Phillips
- **JNT** Jewish New Testament: A Translation of the New Testament That Expresses Its Jewishness
- **KJV** King James Version
- **DKJB** Defined King James Bible
- **KJII** King James Version II (renamed to Literal Translation of the Holy Bible)
- **KJ21** King James for the 21st Century
- **KJ2000** King James 2000
- **LITV** The Literal Translation of the Holy Bible (formerly named King James II)
- **MKJV** Modern King James Version
- **RAV** Revised Authorised Version (British edition of the NKJV)
- **RKJV** Revised King James New Testament
- **TMB** The Third Millennium Bible
- **UKJV** Updated King James Version
- **LB** Living Bible
- **MAEV** Modern American English Vernacular
- **MLB** Modern Language Bible: New Berkeley Version
- **Mof** Bible: James Moffatt Translation
- **NAB** New American Bible
- **NASB** New American Standard Bible
- **MLB** New Berkeley Version (see Modern Language Bible)
- **NCV** New Century Version
- **NEB** New English Bible
- **NET** New English Translation
- **NEvT** New Evangelical Translation
- **NIrV** New Internation Reader's Version
- **NIV** New International Version
- **NJB** New Jerusalem Bible
- **NKJV** New King James Version
- **NLV** New Life Version
- **NLT** New Living Translation
- **NRSV** New Revised Standard Bible
- **NWT** New World Translation (published by the Watchtower Bible and Tract Society of the Jehovah's Witnesses)
- **OBP** The Original Bible Project
- **OSB** Orthodox Study Bible
- **ONT** The Original New Testament: The First Definitive Translation of the New Testament in 2000 Years, by Hugh Schonfield
- **PMB** Postmodern Bible - Amos
- **Rec** Recovery Version
- **REB** The Revised English Bible (revision of NEB)
- **RSV** Revised Standard Version
- **RV** Revised Version, 1885
- **Sch** The Schocken Bible

- **SEB** The Simple English Bible
- **TM** The Message
- **TMB** The Third Millennium Bible
- **TEV** Today's English Version
- **TNIV** Today's New International Version
- **Tyn** Tyndale
- **Wey** Weymouth
- **WEB** World English Bible
- **Wms** The New Testament in the Language of the People, by Charles B. Williams)
- **WNT** Wesley's New Testament
- **Wuest** The New Testament (An Expanded Translation)
- **Wyc** Wycliffe
- **Yes** Yes Word (update of Tyndale translation)
- **YLT** Young's Literal Translation of the Bible

Appendix D.1.2 Non-English Modern Languages

Thousands of additional languages have Bibles or portions; most of these have only one translation in the language. In those cases the language code as defined elsewhere in OSIS may be used, with no following name required.

- **Luther** German by Martin Luther, 1534
- **Algonquin** Tr. John Eliot, 1662
- **ReinaV** Spanish Reina Valera

Appendix E Sample Work Declarations

The following are sample **work** declarations for some of the more common Bibles that will be encountered in actual encoding practice.

Example 1:

```
<work osisWork="thisWork">
  <title>New International Version</title>
  <type type="OSIS">Bible</type>
  <identifier type="OSIS">Bible.en.NIV.1984</identifier>
  <rights type="x-copyright">American Bible Society 1994</rights>
  <refSystem>Bible</refSystem>
</work>
```

Example 2:

```
<work osisWork="kjv">
  <title>King James Version</title>
  <type type="OSIS">Bible</type>
  <identifier type="OSIS">Bible.en.KJV</identifier>
  <rights type="x-copyright">American Bible Society</rights>
  <refSystem>Bible</refSystem>
</work>
```

Appendix F USFM to OSIS Mapping

The mapping table lists the USFM marker first, followed by an XPath that describes how OSIS would encode the same material. Any notes on the USFM code are in the last column. This table made possible by the assistance of SIL.

USFM to OSIS Mapping

USFM Code	OSIS Code	USFM Notes
add	transChange[@type="added"]	add...add* -Special - Translational Addition
b	See notes below on \m and \b	b - Poetry - Stanza Break (Blank Line)
bd	hi[@type="bold"]	bd...bd* - Character - Bold Text
bdit	hi[@type="bold"]/hi[@type="italic"]	bdit...bdit* - Character - BoldItalic Text
bk	reference[@type="x-bookName"]	bk...bk* - Character - Quoted book title
c		
ca	If it is desired to have two "n" attribute values then encode two **chapter** elements - otherwise simply put one or more identifier in the osisID and up to one value for "n".	ca - Chapter Number - Alternate (see notes at right)
cd	div[@type="book"]/div[@type="introduction" and preceeding-sibling::chapter] -- e.g. make this an intro div which immediately follows a chapter milestone	cd - Chapter - Description (introductory text descibing the chapter contents)
cl	title type="chatperLabel"	cl - Chapter - Publishing Label (the text lable for a chapter division heading, e.g. "Chapter")
cls	closer	cls - Paragraph - Closure of an Epistle
conc	div[@type="concordance"]	conc Peripherals-Back Matter Concordance
cov	div[@type="coverPage"]	cov Peripherals - Other, Cover

cp	If it is desired to have two "n" attribute values then encode two **chapter** elements - otherwise simply put one or more identifier in the osisID and up to one value for "n".	cp - Chapter Number - Publishing Alternate (see notes at right)
d	title[@type="psalm"]	d - Heading - Descriptive Title, Hebrew Subtitle
dc	transChange[@type="added" and @edition="dc"]	dc-Special-Deuterocanonical/LXX Additions
em	hi[@type="emphasis"]	em...em* - Character - Emphasized Text
f	note	f...f* - Footnote
fdc	note[not(@type='crossReference')]/seg[@edition="dc"]	fdc...fdc* - Footnote - DC text
fe	note[@placement="end"]	fe...fe* - Endnote (some translations specify certain notes to occur at page bottom, and others which are considered less important to be included at the end of the book)
fig	figure (map attributes to the "|" separated values)	fig...fig*-Auxiliary- Figure/Illustration/Map
fk	catchWord	fk - Footnote - Keyword
fm	This marker should never be found in field texts. It is internal use only in publishing centers.	fm - Footnote - Additional Caller to Previous Note
fq	note/q	fq - Footnote - Quotation or Alternate Rendering (see notes for \fqa)
fqa	note/rdg	fqa - Footnote - Alternate Translation Rendering (Some translations mark-up the quoted text from the existing translation, as well as an alternate translation both with \fq -- others differentiate these two items using \fqa for the alternate. Using \fq for both is more common.)
fr	note/reference[@type="source"]	fr - Footnote - Reference
ft	text within the **note** element, may serve to indicate the end of text of another format marker	ft - Footnote - Text
fv	seg[@type="verseNumber"]	fv...fv* - Footnote - Embedded Verse Number (used when a longer string of

scripture is being included in footnote text, and includes a verse number)

glo	div[@type="glossary"]	glo Peripherals - Back Matter Glossary
h	div[@type='book']/title[@short='value of h marker']	h - File - Header
h1	div[@type='book']/title[@short='value of h marker']	h1 - File - Header
h2	title[@type="runningHead" and placement="inside" and @short='value of h2 marker']	h2 - File - Left Header
h3	title[@type="runningHead" and placement="outside" and @short='value of h3 marker']	h3 - File - Right Header
ib	See notes below on \m and \b	ib - Introduction - Blank Line (This is a format oriented marker -- whitespace, but needs to be preserved round-trip)
id	description[@type='usfm' and @subType='x-id']	id - File - Identification
ide	description[@type='usfm' and @subType='x-ide']	ide - File - Encoding (the document encoding, or the name of a custom font associated with the text)
idx	div[@type="index"]	idx Peripherals - Back Matter Index
ie		ie - Introduction - End Marker (optional)
iex	div[@type="bridge"]	iex - Introduction - Explanatory or Bridge Text
im	See notes below on \m and \b	im - Introduction - Paragraph, no first line indent
imi	Continue the div[@type="introduction" and @canonical="false"]//list/item/p if interrupted other wise div[@type="introduction" and @canonical="false"]//list/item	imi - Introduction - Indented Para, no first line indent
imq	div[@type="introduction" and @canonical="false"]//q	imq - Introduction - Paragraph, quote from text, no first line indent. (This is like \ipq except that the portion quoted is not a whole unit (paragraph) from the text.)
imt	div[@type="introduction" and	imt - Introduction - Major Title Level 1

132

@canonical="false"]/title[@type="main"]

imt1 div[@type="introduction" and @canonical="false"]/title[@type="main"]/title[@level="1"]

imt1 - Introduction - Major Title Level 1

imt2 div[@type="introduction" and @canonical="false"]/title[@type="main"]/title[@level="2"]

imt2 - Introduction - Major Title Level 2

imt3 div[@type="introduction" and @canonical="false"]/title[@type="main"]/title[@level="3"]

imt3 - Introduction - Major Title Level 3

imt4 div[@type="introduction" and @canonical="false"]/title[@type="main"]/title[@level="4"]

imt4 - Introduction - Major Title Level 4

imte div[@type="introduction" and @canonical="false"]/title[@type="main"]

imte - Introduction - [Uncommon] Major Title at Introduction End

intro div[@type="introduction"]

intro Peripherals - Front Matter Introduction

iot div[@type="introduction" and @canonical="false"]//div[@type="outline"]/title

iot - Introduction - Outline Title

io1 div[@type="introduction" and @canonical="false"]//div[@type="outline"]/list/item

io1 - Introduction - Outline Level 1

io2 div[@type="introduction" and @canonical="false"]//div[@type="outline"]/list/item/list/item

io2 - Introduction - Outline Level 2

io3 div[@type="introduction" and @canonical="false"]//div[@type="outline"]/list/item/list/item/list/item

io3 - Introduction - Outline Level 3

io4 div[@type="introduction" and @canonical="false"]//div[@type="outline"]/list/item/list/item/list/item/list/item

io4 - Introduction - Outline Level 4

ior div[@type="introduction" and @canonical="false"]//div[@type="outline"]/list/item/reference

ior...ior* - Introduction - Outline References Range (the scope of the previous outline entry)

ip div[@type="introduction" and @canonical="false"]//p

ip - Introduction - Paragraph

ipi	div[@type="introduction" and @canonical="false"]//list/item/p	ipi - Introduction - Indented Para, first line indent
ipq	div[@type="introduction" and @canonical="false"]//q/p	ipq - Introduction - Paragraph, quote from text. (This is a quote of a portion of scripture in the introduction)
ipr	div[@type="introduction" and @canonical="false"]//reference	
iq	div[@type="introduction" and @canonical="false"]//lg/l[@level="1"]	iq - Introduction - Poetry Level 1
iq1	div[@type="introduction" and @canonical="false"]//lg/l[@level="1"]	iq1 - Introduction - Poetry Level 1
iq2	div[@type="introduction" and @canonical="false"]//lg/l[@level="2"]	iq2 - Introduction - Poetry Level 2
iq3	div[@type="introduction" and @canonical="false"]//lg/l[@level="3"]	iq3 - Introduction - Poetry Level 3
is	div[@type="introduction" and @canonical="false"]/div[@type="section"]/title	is - Introduction - Section Heading Level 1
is1	div[@type="introduction" and @canonical="false"]/div[@type="section"]/title	is1 - Introduction - Section Heading Level 1
is2	div[@type="introduction" and @canonical="false"]/div[@type="section"]/div[@type="subSection"]/title	is2 - Introduction - Section Heading Level 2
it	hi[@type="italic"]	it...it* - Character - Italic Text
k	seg[@type="keyword"]	k...k* - Character - Keyword
li	list/item	li - List Entry - Level 1
li1	list/item	li1 - List Entry - Level 1
li2	list/item/list/item	li2 - List Entry - Level 2
li3	list/item/list/item/list/item	li3 - List Entry - Level 3
li4	list/item/list/item/list/item/list/item	li4 - List Entry - Level 4
lit	lg[@type='doxology']/l[@type='refrain']	lit - Special - Liturgical note (indication fo

134

		what should take place, or be spoken by congregation etc. at a specific point in the text, e.g. \lit Glory!)
m	See notes below on \m and \b	m - Paragraph - Margin, No First Line Indent (usually a continuation of paragraph text which existed before an interrupting element (embedded text) -- but not always).
maps	div[@type="map"]	maps Peripherals - Back Matter Map Index
mi	See notes below on \m and \b	mi - Paragraph - Indented, No First Line Indent
mr	div[@type="majorSection"]/title[@type="scope"]/reference	mr - Heading - Major Section References
mt	div[@type="book"]/title[@type="main"]	mt - Title - Major Title Level 1
mt1	div[@type="book"]/title[@type="main"]/title[@level="1"]	mt1 - Title - Major Title Level 1
mt2	div[@type="book"]/title[@type="main"]/title[@level="2"]	mt2 - Title - Major Title Level 2
mt3	div[@type="book"]/title[@type="main"]/title[@level="3"]	mt3 - Title - Major Title Level 3
mt4	div[@type="book"]/title[@type="main"]/title[@level="4"]	mt4 - Title - Major Title level 4
mte	div[@type="book"]/title[@type="main"]	mte - Title - [Uncommon] Major Title Ending Level 1
mte1	div[@type="book"]/title[@type="main"]/title[@level="1"]	mte1 - Title - [Uncommon] Major Title Ending Level 1
mte2	div[@type="book"]/title[@type="main"]/title[@level="2"]	mte2 - Title - [Uncommon] Major Title Ending Level 2
ms	div[@type="majorSection"]/title	ms - Heading - Major Section Level 1
ms1	div[@type="majorSection"]/title	ms1 - Heading - Major Section Level 1
ms2	div[@type="majorSection"]/div[@type="majorSection"	ms2 - Heading - Major Section Level 2

]/title

nb		nb - Paragraph - No Break with Previous Paragraph (used when a paragraph spans a chapter boundary -- Paratext requires that a paragraph marker be included after a chapter boundary)
nd	divineName	nd...nd* - Character - Name of Deity
ndx	index[@name="subject",@level1="..."]	ndx...ndx* - Auxiliary - Subject Index Entry
no	hi[@type="normal"]	no...no* - Character - Normal Text
ord	hi[@type="super"]	ord...ord* - Character - Ordinal number text portion (the "st" part of "1st")
p	p	p - Paragraph - Normal, First Line Indent
pb	milestone[@type="pb"]	pb - Break - Page Break
pc	inscription	pc - Paragraph - Centered
ph	list/item	ph - Paragraph - Hanging Indent, Level 1 (deprecated, but still needing conversion)
ph1	list/item	ph1 - Paragraph - Hanging Indent, Level 1 (deprecated, but still needing conversion)
ph2	list/item/list/item	ph2 - Paragraph - Hanging Indent, Level 2 (deprecated)
ph3	list/item/list/item/list/item	ph3 - Paragraph - Hanging Indent, Level 3 (deprecated, but still needing conversion)
phi	treat same as ph	DEPRECATED phi - Paragraph - Indented, Hanging Indent
pi	list/item/p	pi - Paragraph - Indented, Level 1, First Line Indent
pi1	list/item/p	pi1 - Paragraph - Indented, Level 1, First Line Indent
pi2	list/item/list/item/p	pi2 - Paragraph - Indented, Level 2, First Line Indent

pi3	list/item/list/item/list/item/p	pi3 - Paragraph - Indented, Level 3, First Line Indent
pmo	q[@type="embedded"]/salute	pmo - Paragraph - Embedded Text Opening
pm	q[@type="embedded"]/p	pm - Paragraph - Embedded Text
pmc	q[@type="embedded"]/closer	pmc - Paragraph - Embedded Text Closing
pmr	q[@type="embedded"]/p[@type="x-refrain"]	pmr - Paragraph - Embedded Text Refrain
pn	name	pn...pn* - Character - Proper Name
pr	p[@type="x-refrain"]	pr - Paragraph - Right Aligned (nearly always used for a refrain text, or for end of quote e.g. "says the LORD") same as pmr
pref	div[@type="preface"]	pref Peripherals - Front Matter Preface
ps	p	DEPRECATED ps - Paragraph - No Break with Next Paragraph
psi	treat same as pi	DEPRECATED psi - Paragraph - Indented, No Break with Next
pub	div[@type="publicationData"]	pub Peripherals - Front Matter Publication Data
pubi nfo	Encode in using the Doublin Core elements in **work** in the header. If for presentation purposes this may be encoded within div[@type="publicationData"]	pubinfo - Publication - Information
q	lg/l[@level="1"]	q - Poetry - Indent Level 1, Single Level Only
q1	lg/l[@level="1"]	q1 - Poetry - Indent Level 1
q2	lg/l[@level="2"]	q2 - Poetry - Indent Level 2
q3	lg/l[@level="3"]	q3 - Poetry - Indent Level 3
qa	lg/l[@type="acrostic"]	qa - Poetry - Acrostic Heading/Marker
qac	hl[@type="acrostic"]	qac...qac* - Poetry Text - Acrostic Letter (to highlight the first letter of a word within lines of text, to show acrostic structure)

qc	l[@type="doxology"]	qc - Poetry - Centered
qm	q[@type="embedded"]/lg/l[@level="1"] (or leave off the level)	qe - Poetry - Embedded Text, Indent Level 1, Single Level Only
qm1	q[@type="embedded"]/lg/l[@level="1"]	qe1 - Poetry - Embedded Text, Indent Level 1
qm2	q[@type="embedded"]/lg/l[@level="2"]	qe2 - Poetry - Embedded Text, Indent Level 2
qm3	q[@type="embedded"]/lg/l[@level="3"]	qe3 - Poetry - Embedded Text, Indent Level 3
qr	l[@type="refrain"] and l[@type="attribution"] (choose based on context)	qr - Poetry - Right Aligned
qs	lg/l[@type="selah"]	qs...qs* - Poetry Text - Selah
qt	seg[@type="otPassage"]	qt...qt* - Special - Quoted Text, OT in NT
r	title[@type='parallel']/reference	r - Heading - Parallel References
rem	description[@type="usfm" and subType="x-rem"]	rem - File - Remark
resto re	description[@type="usfm" and subType="x-restore"]	restore - File - Restore Information (a code inserted by Paratext when restoring a project)
s	div[@type="section"]/title	s - Heading - Section Level 1
s1	div[@type="section"]/title	s1 - Heading - Section Level 1
s2	div[@type="subSection"]/title	s2 - Heading - Section Level 2
s3	div[@type="subSection"]/div[@type="subSection"]/title	s3 - Heading - Section Level 3
s4	div[@type="subSection"]/div[@type="subSection"]/div[@type="subSection"]/title	s4 - Heading - Section Level 4
sc	hi[@type="small-caps"]	sc...sc* - Character - Small Caps
sig	signed	sig...sig* - Character - Author's Signature (Epistles)

sls	foreign[@type="x-secondaryLanguage"]	sls...sls* - Special - Secondary Language Source (to specify where the original source language is different that the primary source -- e.g. Aramaic source in the OT)
sp	speech/speaker	sp - Heading - Speaker
spine	div[@type="spine"]	spine Peripherals - Other, Spine
tc1	row/cell	tc1 - Table - Column 1 Cell
tc2	row/cell	tc2 - Table - Column 2 Cell
tc3	row/cell	tc3 - Table - Column 3 Cell
tc4	row/cell	tc4 - Table - Column 4 Cell
tcr1	row/cell[@align='right']	tcr1 - Table - Column 1 Cell, Right Aligned
tcr2	row/cell[@align='right']	tcr2 - Table - Column 2 Cell, Right Aligned
tcr3	row/cell[@align='right']	tcr3 - Table - Column 3 Cell, Right Aligned
tcr4	row/cell[@align='right']	tcr4 - Table - Column 4 Cell, Right Aligned
th1	row[@role="label"]/cell	th1 - Table - Column 1 Heading
th2	row[@role="label"]/cell	th2 - Table - Column 2 Heading
th3	row[@role="label"]/cell	th3 - Table - Column 3 Heading
th4	row[@role="label"]/cell	th4 - Table - Column 4 Heading
thr1	row[@role="label"]/cell[@align='right']	thr1 - Table - Column 1 Heading, Right Aligned
thr2	row[@role="label"]/cell[@align='right']	thr2 - Table - Column 2 Heading, Right Aligned
thr3	row[@role="label"]/cell[@align='right']	thr3 - Table - Column 3 Heading, Right Aligned
thr4	row[@role="label"]/cell[@align='right']	thr4 - Table - Column 4 Heading, Right Aligned

tl	foreign	tl...tl* - Character - Transliterated Word
toc	div[@type="tableOfContents"]	toc Peripherals - Front Matter Table of Contents
tr	table/row	tr - Table - Row
tr1	treat same as tr	DEPRECATED tr1 - Table - Row, Level 1
tr2	treat same as tr	DEPRECATED tr2 - Table - Row, Level 2
v	verse\@osisID and verse\@n	v - Verse Number
va	If it is desired to have two "n" attribute values then encode two **verse** elements - otherwise simply put one or more identifier in the osisID and up to one value for "n".	va - Verse Number - Alternate
vp	If it is desired to have two "n" attribute values then encode two **verse** elements - otherwise simply put one or more identifier in the osisID and up to one value for "n".	vp - Verse Number - Publishing Alternate
w	index[@level1="..."]...	w...w* - Auxiliary - Wordlist Entry
wh	index[@name="hebrew",@level1="..."]...	wh...wh* - Auxiliary - Hebrew Wordlist Entry
wg	index[@name="greek",@level1="..."]...	wg...wg* - Auxiliary - Greek Wordlist Entry
x	note[@type="crossReference"]	x...x* - Cross Reference
xo		
xt	text within the **note** element, may serve to indicate the end of text of another format marker	xt - Cross Reference - Target References
xk	catchWord	xk - Cross Reference - Keyword
xq	q	xq - Cross Reference - Quotation
xdc	note[@type='crossReference']/seg[@edition="dc"]	xdc...xdc* - Cross Reference - DC Target Refs (cross reference which should only be rendered for editions which include the DC books)

Appendix F.1 Notes on \b, \m, and \mi

If \b immediately follows a line group, list, or table and is not immediatelly followed by \s, \s#, \p, \ms, \ms#, or \ps, then the \b is simply indicating that the line group, list, or table is closing, OTHERWISE an explicit line break must be encoded

\m immediately following a line group, list, or table closes the line group, list, or table and continues the text that follows in the already open container including a **div** element OTHERWISE an explicit line break must be inserted and the text that follows in the already open container including **div** (if a \m immediately follows a \b then an explicit line break is inserted)

\mi immediately following a line group, list, or table closes the line group, list, or table and continues the text that follows in the parent of the line group, list, or table if an ancestor is q[@type="embeded] OTHERWISE a q[@type='embedded'] is inserted into the parent of of the line group, list, or table, OTHERWISE if a q[type='embedded'] is an ancestor insert an explicit line break. If not q[@type='embedded'] is an ancestor then open one and continue the text that follows with no line break.

if the current container or one of its ancestors is a q[type='embedded'] then insert a explicit line break and continue within the current container OTHERWISE open a q[@type='embeded'] and continue within it

\ib, \im, and \imi behave the same as \b, \m, and \mi but are within a div[@type='introduction']

In addation and overriding the above, if a line break (
) element would be encoded as the first node (element node or text node) within a **div**, **p**, or **q** element so as to represent a \m or \mi concept, the line break element is unnecessary and should be ommited.

In all of the above rules regarding \b, \m, \mi, \imi, and \im a **note** is ignored when determining immediately following, immediately preceding, and first relationships.

Appendix G USMARC Relator Codes

For ease of use a selected set of USMARC Relator Codes have been extracted from the larger complete list of such codes. This short listing should be sufficient for basic encoding needs. The complete list of USMARC Relator Codes can be found at: http://lcweb.loc.gov/marc/relators/re0002r1.html.

Appendix G.1 Selected Contributor Roles

- **ann** Annotator: Use for a person who writes manuscript annotations on a printed item.
- **art** Artist: Use for a person (e.g., a painter) who conceives, and perhaps also implements, an original graphic design or work of art.
- **aut** Author: Use for a person or corporate body chiefly responsible for the intellectual or artistic content of a work, usually printed text. This term may also be used when more than one person or body bears such responsibility.
- **cwt** Commentator for written text: Use for a person or corporate body responsible for the commentary or explanatory notes about a text. For the writer of manuscript annotations in a printed book, use Annotator.
- **com** Compiler: Use for a person who produces a work or publication by selecting and putting together material from the works of various persons or bodies.
- **ctb** Contributor: Use for one whose work has been contributed to a larger work, such as an anthology, serial publication, or other compilation of individual works. Do not use for someone whose sole function in relation to a work is as author, editor, compiler or translator.
- **cre** Creator: Use for a person or corporate body responsible for the intellectual or artistic content of a work.
- **edt** Editor: Use for a person who prepares for publication a work not primarily his/her own, such as by elucidating text, adding introductory or other critical matter, or technically directing an editorial staff.
- **ill** Illustrator: Use for the person who conceives, and perhaps also implements, a design or illustration, usually to accompany a written text.
- **pbl** Publisher
- **trl** Translator: Use for a person who renders a text from one language into another, or from an older form of a language into the modern form.

Appendix H Identifying a Work given a work declaration element

The elements **identifier** type = OSIS is the primary means of identifying a referenced work.

Appendix I Special Case: Two References Systems

It is unlikely that most encoders will ever encounter a text with two references systems, much less ones that do not start and end at the same points. Unless you are facing that situation or just like difficult explanations, please skip this section.

Can't say you were not given fair warning! There are three types of cases that are likely to arise when there are two (or more) reference systems in use with one text.

- The two (or more) reference systems share start/end points, or
- Where one reference system is a subset of the other reference system, or
- The two (or more) reference system have different start/end points.

The simpler case, where start and end points coincide is treated first, followed by the more complex one with differing start and end points (also known as overlapping).

In the case were two reference systems share start and end points, the OSIS best practice solution is to encode the reference value of each system in the osisID for each element. That is to say that a verse, for example, would have an osisID just like it would for any other text, and, it will have a second value in the osisID that has a work prefix. That is to say there is no defaulting of the work prefix on the second osisID value. An example of such an encoding is:

```
<p><verse sID="A"  osisID="Ps.1.1 h:Ps.1.2"/>
text text texttext text text
<verse eID="A"/><verse sID="B" osisID="Ps.1.2 h:Ps.1.3"/>
text text text text<verse eID="B"/></p>
```

Note that the material that begins with "h:" is the second reference system and the prefix, must be a reference to a work found in the header of the document. (Last chance to bail before we get to the really ugly one.)

A more complex case is presented when the two reference systems do not agree where elements start and end. While alternatives do exist, the OSIS method for dealing with this problem is called milestones. A milestone, like the ones along side a highway, simply mark a location in a particular spot. OSIS milestones, also tell you if they are the starting or ending milestone, plus carry other attributes associated with the element they are representing.

Using OSIS milestones, the case where two reference systems do not agree on start and end points can be shown as:

1. <p>
2. <verse sID="A" osisID="Ps.1.1"/>text text <verse eID="0"/>
3. <verse sID="1" osisID="h:Ps.1.1"/>text text
4. <verse eID="1"/>
5. <verse sID="2" osisID="h:Ps.1.2"/>text text
6. <verse eID="A"/>
7. <verse sID="B" osisID="Ps.1.2"/>text text
8. <verse eID="2"/>
9. <verse sID="3" osisID="h:Ps.1.3"/>
10.<verse eID="B"/>
11.</p>

While the display may look rather confusing, it actually works quite simply. Start at line 2, for example. The osisID is **Ps.1.1** with no prefix. That indicates that this is in the declared default reference system for this work. Now look down the list until you see an attribute called **eID** on a verse that has a value that matches that of line 2. That's correct, line six has a verse with eID="A" so it matches the verse at the beginning of line 2.

In full the matching is as follows:

1. matches line 11

2. matches line 6
3. matches line 4
4. matches line 3
5. matches line 8
6. matches line 2
7. matches line 10
8. matches line 6
9. matches a line not shown (later)
10. matches line 7
11. matches line 1

Appendix J osisIDs: Construction Rules

The really adventurous reader will consult the osisCore.2.1 schema for the regular expression that governs the form of osisIDs. For those in a hurry or who simply want to avoid the complexity of XML Schema regexes (the abbreviated form of regular expressions) the following guide should suffice.

Any **osisID** is divided into a number of parts, some of which are optional, that is they can be omitted and still have a valid **osisID**. The following breaks out the structure of an **osisID** into its various parts and notes what is allowed in each part and what parts are required.

Appendix J.1 Prefix: (optional)

The prefix to an **osisID** must contain at least one letter, number or underscore, that may be followed by any number of letters, numbers or underscores, separated by periods, and concluding in a colon ":". Note that if you use a prefix, the colon is required. The prefix is **optional**.

Some examples of valid prefixes include:

- Bible:
- Bible.French:
- Spurgeon.Commentaries_Job:

Appendix J.2 Main (required)

The main part of an **osisID** consists of at least one letter, number or underscore, that may be followed by any number of letters, numbers or underscores, separated by periods. The main part of the **osisID** is **required**.

Note that one difference from traditional identification of Bible verses that OSIS uses a period to separate the verse from the chapter. One usually sees, Gen. 1:1. That is what should be displayed to the reader of a OSIS text, but use of whitespace as a separator (between Gen. and chapter 1, in XML causes problems. So, the whitespace was replaced by a period.

Some examples of valid main parts of an osisID include:

- Gen.
- Mark
- Mark.8
- Matt.6.1

Appendix J.3 Extension (optional)

While standard citation systems are well known and should be covered by the main part of the osisID, there are cases where such systems have been extended. Some of those extensions are standard and others are not.

In order to allow for extension of citation systems, OSIS allows a standard citation to be followed by the exclamation mark "!" which signals that what follows is not part of the standard reference. This allows

systems that do not recognize extensions to at least put the user at the starting place of the standard reference.

The beginning exclamation mark is required, if the extension mechanism is used and is followed by least one letter, number or underscore, that may be followed by any number of letters, numbers or underscores, separated by periods.

Some examples of valid extensions to an **osisID** include:

- Prov.26.12!b ID for the second half of verse 12.
- other examples?

Appendix K osisRefs: Construction Rules

The **osisRef** regex is over twice as long as the **osisID** regex, in part because of the additional capabilities of an **osisRef**. The allowable characters are basically the same but there are some nuances to constructing an **osisRef**. The following guide should get you past all of the common cases, and even a few of the odder ones.

Appendix K.1 Prefix: (optional)

The prefix to an **osisRef** must contain at least one letter, number or underscore, that may be followed by any number of letters, numbers or underscores, separated by periods, and concluding in a colon ":". Note that if you use a prefix, the colon is required. The prefix is **optional**.

Note that if you omit the prefix on an **osisRef**, it is optional afterall, your reference can only point to another location in the OSIS text where you are inserting the **osisRef**. This is the equivalent of the **osisID** without a prefix, it defaults to the text that you are working in at the moment. For purposed of illustration, all the **osisRef**s shown below have the prefix attached.

Some examples of valid prefixes include:

- Bible:
- Bible.French:
- Spurgeon.Commentaries_Job:

Appendix K.2 Main (required)

The main part of an **osisRef** consists of at least one letter, number or underscore, that may be followed by any number of letters, numbers or underscores, separated by periods. The main part of the **osisRef** is **required**.

Note that one difference from traditional identification of Bible verses that OSIS uses a period to separate the verse from the chapter. One usually sees, Gen. 1:1. That is what should be displayed to the reader of a OSIS text, but use of whitespace as a separator (between Gen. and chapter 1, in XML causes problems. So, the whitespace was replaced by a period.

Some examples of valid main parts of an **osisRef** include:

- Gen.
- Mark
- Mark.8
- Matt.6.1

Appendix K.3 Extension (optional)

While standard reference systems are well known and should be covered by the main part of the **osisRef**, there are cases where such systems have been extended. Some of those extensions are standard and other are not.

In order to allow for references that use an extension of citation systems, OSIS allows a standard citation to be followed by the exclamation mark "!" which signals that what follows is not part of the standard reference. This allows systems that do not recognize extensions to at least put the user at the starting place of the standard reference and avoids having to define a reference system for most works.

The beginning exclamation mark is required, if the extension mechanism is used and is followed by least one letter, number or underscore, that may be followed by any number of letters, numbers or underscores, separated by periods.

Some examples of valid extensions to an **osisRef** include:

- Prov.26.12!b osisRef for the second half of verse 12.
- other examples?

Appendix K.4 Grains (optional)

One shortcoming of most reference systems is the inability to point to a particular place in a line of text. This is of particular interest for Bible study, where the user wants to point to a particular word in a passage, not the entire passage itself. OSIS developed a syntax that follows the prefix, main osisRef and even the extension (if present) that allows you to do exactly that.

The grain operators come in two types: 1) **cp**, which allows you to point at a particular character in the text, and 2) **s**, which allows you to point at a string of characters. It is probably easier to illustrate these separately.

The **cp** grain operator is a number, enclosed by square brackets and preceded by the "@" sign, all of which follows, at a minimum, the main part of an **osisRef**. For example:

- RSV:Gen.1.1@cp[8] Points at the starting character of the word "beginning."
- RSV:Gen.3.20@cp[32] Points to the starting character of the word "Eve."

This operator will be the most useful for automated systems that allow users to point and select a point in the text for automatic generation of this operator. When this syntax was being developed, the editors made the mistake of picking an example before considering how tedious it was to count spaces, apostrophes, and other punctuation that goes into the total for a **cp** operator. Users who wish to avoid the tedium of (and error prone as well) counting characters, may wish to use the **s** operator.

The **s** grain operator is a string, enclosed by square brackets and preceded by the "@" sign, all of which follows, at a mimimum, the main part of an **osisRef**. For example:

- RSV:Gen.1.1@s[beginning] Points at the starting character of the word "beginning."
- RSV:Gen.3.20@s[Eve] Points to the starting character of the word "Eve."

You may wish to convince yourself that the **s** operator is easier to use than **cp** but to each his own.

Warning: Note that the **s** operator does not allow spaces. That is to say that you cannot put a phrase between the square brackets. That limitation is due to the handling of spaces in XML. It was an issue that the editors struggled with for some time but ultimately, it was decided that word level matching would meet most users needs.

Appendix K.5 Ranges (optional)

It is often the case that texts make references to a range of Bible verses and with the osisRef mechanism, not only duplicates that ability, but also provides for the grain matching mentioned above.

The beginning of a range in an osisRef is indicated by a hyphen "-" character that occurs at the very end of the first part of the range. That hyphen is immediately followed by the same order of expression found in the first part, with one exception, there is no prefix allowed on the second half of an **osisRef** range.

The reason to disallow a prefix on the second half of a range is quite simple. A range, at least in the OSIS sense, is defined as occurring within a work. That is to say that a range that attempted to say: 'Bible:Gen.1.1-Livy:Bk.1', would make no sense to any processor. So, when using the range operator, be sure that the range occurs within a single work.

With the omission of the prefix, the second half of a range follows the same rules as the first half.

Appendix L Conformance requirements

Appendix L.1 Conformance levels

There are 4 levels of OSIS conformance for the markup in OSIS documents:

Appendix L.1.1 Level 1: "Minimal OSIS document"

The document must be a well-formed and valid XML document according to the OSIS schema.

The document must be complete in accordance with the scope declaration in its work declaration. For example, a document with a missing chapter is not OSIS-conforming.

The document must mark all canonical references where applicable (for example, book, chapter, and verse boundaries in Bibles. Marking in groups, for example a paragraph that includes several verses, is permissible.

The header must include work declarations for the document itself, and for the versification system it uses.

All work declarations must provide unique **osisWorkID** values, and **only** those values may appear as work identifiers in osisIDs and osisRefs (whether by default or explicit) in the document.

All work declarations must provide at least title, creator, and date(s). Creator may be coded as "(anonymous)" or "(unknown)" if applicable. The date of electronic publication is required; other dates may be omitted or coded as "(unknown)" if applicable, though they should be provided if known.

At least one revision description element must be included, describing the most recent substantial changes to the document. The name and email address of the last responsible party should be included.

Empty elements substituted for containers (such as verse, q, etc.) must occur in matched pairs. Each end must actually be expressed by a true XML empty element, not by start and end tags with nothing between. The earlier member of each pair must have an sID attribute and no eID attribute; the later member of each pair must have an eID attribute and no sID or any other attributes. The sID and eID values for a pair must match (including as to case), and must be distinct from all other sID and eID attribute values in the document.

All elements must be used substantially in accordance with their intended meaning as conveyed in this documentation (including documentation and standards referred to, such as Dublin Core, USMARC Relator Codes, and so on).

Appendix L.1.2 Level 2: "Basic OSIS Document"

All requirements of Level 1 conformance must be fulfilled.

A clear statement of rights must be provided within the **rights** element. If the document is licensed for free copying under certain conditions, those conditions or a reliable URI to them must be provided. If there are encumbrances or if clearance is required to copy or use the work, contact information for the responsible party must be provided directly within the **rights** element.

The source edition from which the electronic edition was produced must be clearly identified, or clearly stated as unknown (the latter practice is deprecated, and encoders are strongly encouraged to make a serious effort to identify the source edition).

All inscriptions (for example, "mene mene tekel parsin") must be marked where applicable.

All instances, translations, or transliterations of the tetragrammaton must be marked via the **divineName** element.

All languages substantially appearing in the text must be identified, and all points where the text itself identifies a phrase as coming from a particular language must be marked up to match (for example, "Talitha cumi").

All epistolary markup (opener, closer, signature, salute) must be provided where applicable.

Poetic text must be marked sufficiently to enable rendering it readably as poetry. The distinction of using **l** for linguistically or poetically significant line breaks, versus using **lb** for typographically significant or preferred line breaks, must be maintained.

If the source edition had section, paragraph, block quotation, or other similar demarcations in addition to book, chapter, and verse numbering, they must be included and appropriately marked up.

If the source edition had footnotes, sidenotes, endnotes, or other notes, they must be included, and must be distinguished into as many types as can be readily distinguished by observing the typographic conventions of the source edition. Once OSIS standardizes a format for external annotation files, this requirement may be fulfilled either by inline encoding of annotations, or external.

Appendix L.1.3 Level 3: Complete OSIS document

All the requirements of Level 2 must be fulfilled.

All notes, front and back matter, illustrations, section heads, and other non-canonical phenomena of the source edition must be included.

Appendix L.1.4 Level 4: Scholarly OSIS document

All the requirements of Level 3 must be fulfilled.

Substantial critical apparatus must be available in the text, such as: Strong's or comparable numbering of words; part-of-speech and/or other linguistic markup; encoding of variant readings, critical apparatus, and the like; extensive translation, scholarly, interpretive, or other notes.

At least highly significant persons and places in the text must be marked as names, and refer to the normative form of the corresponding individual (the Bible Technologies Group is preparing normative lists at this time). Where such identification is a matter of non-obvious interpretation, that fact must be marked, and the encoders' practices and biases should be duly noted in the front matter.

The text must also conform to the requirements of Level 3 Quality as described below.

Appendix L.2 Quality levels

The conformance levels defined above do not specify the level of accuracy and proofreading of the text proper. This is instead measured by the following scale of "Quality":

Appendix L.3 Level 1: Sub-OCR Quality

The text may have many typographical errors; essentially, it is unproofed text from automated OCR, probably of a less-than-ideal original.

Appendix L.4 Level 2: OCR Quality

The text may have up to 5 typographical errors per source page. It may be unproofed output from ideal OCR of an ideal source, or may have been run at least through rudimentary spell-checking or vocabulary counting and repair, or entered by a double-keying or similar service that maintains accuracy to the required level.

Appendix L.5 Level 3: Proof Quality

There may not be more than an average of 1 error per source page (or per 2000 characters of content) as compared with the stated copy text. This requirement does not preclude producing new editions, which for example may fix typos in the original, normalize spelling of older texts, and so on. However, in such cases it is recommended that the best available copy of the source text as it existed prior to such modernizations, also be made available.

Appendix L.6 Level 4: Trusted Quality

A Trusted Quality document must fulfill all the requirements of a Proof Quality document, and must also have been in public use for at least one year, and read by at least 5 independent proofreaders, with all noted errors fixed. The text should have available a complete log of changes made since it reached Proof Quality. Random spot-checks of at least 3% of the text must come up with no instances of more than 1 error per 5 pages (or 10,000 characters of content).

Appendix M Application Requirements

Applications should avoid making any processing distinctions between elements represented as non-crossing single elements or as milestone pairs.

Applications must interpret OSIS references as accurately as is feasible, but apply smart fallback as needed. For example, grains will not map across translations or languages, though most will typically survive changes between successive editions of the same text, or differences between British and American English versions. Applications should in general at least offer to take the user to the nearest reliably-findable place; in this case, the verse.

Applications must be able to interpret the OSIS elements and process them in a manner consistent with their express intent as specified in this document, and in accordance with standard practices of Bible publishing. For example, applications should be capable of distinguishing the typography used for inscriptions, the divine Name, verse labels and references, foreign insertions in the text, notes, and so on in ways readily recognizable to users of print Bibles.

The Bible Technologies group also strongly advocates making all software, and especially all OSIS-aware software, accessible to print-disabled users. This includes details such as providing text alternates for all graphics, not marking up poetry such that it can only be line-broken given certain line widths or font sizes; not making crucial distinctions only via color, subtleties of font, etc.; and not using tables gratuitously to achieve formatting goals rather than to represent truly tabular information. Subtle technical factors can also ruin otherwise accessible software, for example, the order in which panes are drawn. Implementers are strongly encouraged to consult with experts on accessibility, and obtain specific critical testing and review by print-disabled users before finalizing product releases. The Bible Technology will, as resources permit, be glad to help connect implementors with accessibility experts.

Appendix N The Bible Technology Group

BTG is a joint effort that has been supported most tangibly by the American Bible Society and the Society for Biblical Literature, as well as by the United Bible Societies, numerous national Bible Societies, the Summer Institute of Linguistics, and othe organizations.

The work of BTG has also been greatly enhanced by many other members of these and other organizations, who have responded to drafts, made numerous, useful, and sometimes essential recommendations, and encoded texts to test the schemas for usability, consistency, and other virtues. Especially notable among these have been Robin Cover, Jonathan Robie, and Bob Pritchett.

The official Website for BTG is http://www.bibletechnologies.net, and much additional information can be found there.

Appendix O Deprecated Elements

Users with XML aware software will note the appearance of **milestoneStart** and **milestoneEnd** elements wherever the **milestone** element is permitted. Those elements were retained in the OSIS schema for backwards compatibility reasons only. They should **not be used** under any circumstances. The OSIS schema has replaced those elements with the much easier and more robust milestoneable mechanism.

Be aware that these elements may be removed without warning from future editions of the OSIS schema, once all (or most all) users have transitioned to the later mechanisms.