Comparing Tagging of a Table of Contents
Adobe Acrobat Pro, Foxit Phantom for Business
Nuance PowerPDF

Karen McCall, M.Ed.
Copyright 2019

Table of Contents

Documents used for this Tutorial 3
The Accessible Word Document 3
Tables of Content in the Past 3
Adobe Acrobat Pro DC and Microsoft Tagging Tools 5
Foxit Phantom for Business 7
Nuance PowerPDF 8
Summary 10
Appendix A: Contact Information 12

Documents used for this Tutorial

There are two versions of this document, one for general public consumption (no attachments) and one with the attached test files. The one with the attached test files is only available if you purchase the book Accessible and Usable PDF Documents: Techniques for Document Authors, fourth Edition[1].

At this point in time, June 2019, NONE of the applications used for this test gives you a clean, easy to navigate and/or remediate Table of Contents (TOC). Parts of the< TOCI> Tags are being truncated, PAC 3 is demanding that the <Link> Tag have Alt text when it doesn't need Alt Text because it is a <TOCI>, sometimes you get a <Reference> Tag nested in the <TOCI> Tag…

This is the reason I chose my photo of the Cave of the Mounds near Madison Wisconsin for this series…beautiful to look at but could also represent a melting mess!

The Accessible Word Document

I've created a Word document that represents an accessible document with all of the common document elements used on a daily basis. Instead of a Text Box, I've demonstrated that using a Paragraph Style is 100 percent more accessible and offers both consistency of look and feel and more visual display options/formatting. It demonstrates the use of columns instead of a two cell table and using the List Paragraph Style to add space between list items instead of pressing the Enter key.

For the Table of Contents in this demonstration document I used Headings in my Word document and, once I had a few Headings, went to the References Ribbon, Table of Contents and chose Custom/Create Table of Contents. I did not use the Tables of Contents samples from the Table of Contents Gallery as this creates a Table of Contents using Content Controls which are still inaccessible.

The end point is that I even created the most accessible Table of Contents I could. Any cosmetic changes to the default Table of Contents were made using the TOC Styles. I've never had a problem with this before and I've been using this type of document for years in my training and tutorials. In June 2019, I started from scratch and recreated it in Word 2016/Office 365 subscription version just in case there were any legacy bits under the hood. This is a clean document with no legacy Word version stuff in it.

Tables of Content in the Past

Prior to recent updates of Adobe acrobat Pro DC, we saw a sort of uniform tagging of a Table of Contents coming from Word. If you looked at the Tags, they sort of made sense. If

[1] Accessible and Usable PDF Documents: Techniques for Document Authors, Forth Edition available through Teachable: https://karen-mccall.teachable.com/p/accessible-and-usable-pdf-documents

you used a screen reader or Text-to-Speech tool, you got a complete <TOCI> listing which meant the Heading used, the Dot Leaders and the page numbers.

The one annoying thing, even then, was that both the Microsoft tagging tool and the Adobe acrobat Ribbon conversion tool included the words “Table of Contents” or “Contents” in the TOC as a Table of Contents Item (TOCI) EVEN though they were identified as a Subtitle Style in the original Word document and should be separate from the actual TOC!

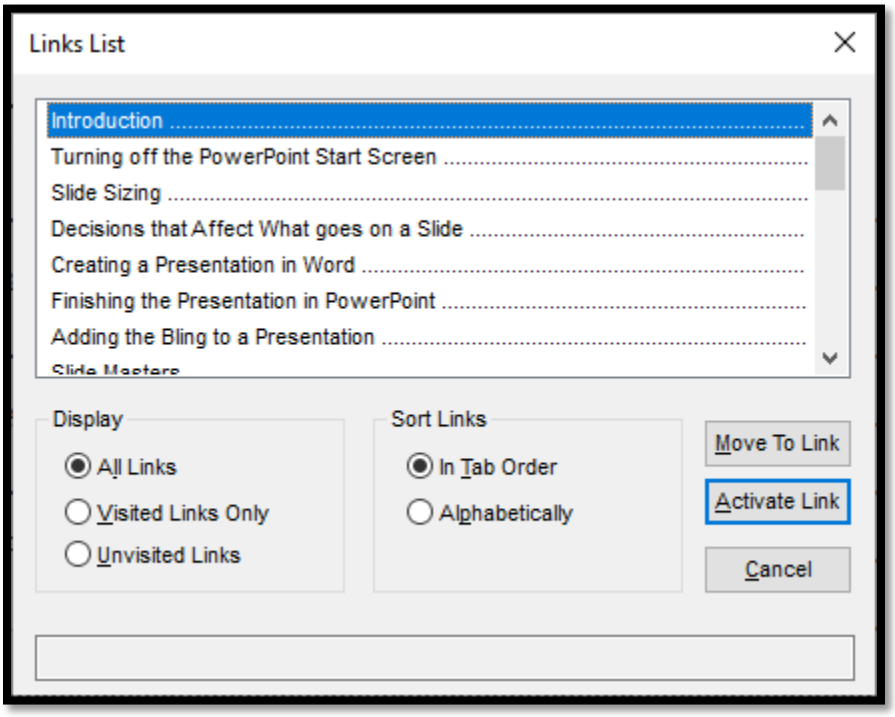

Figure 1 What a correctly tagged Table of Contents should look like to adaptive technology.

The preceding image is from a Word document I wrote on the basics of accessible PowerPoint, published in October 2018…not even a year ago.

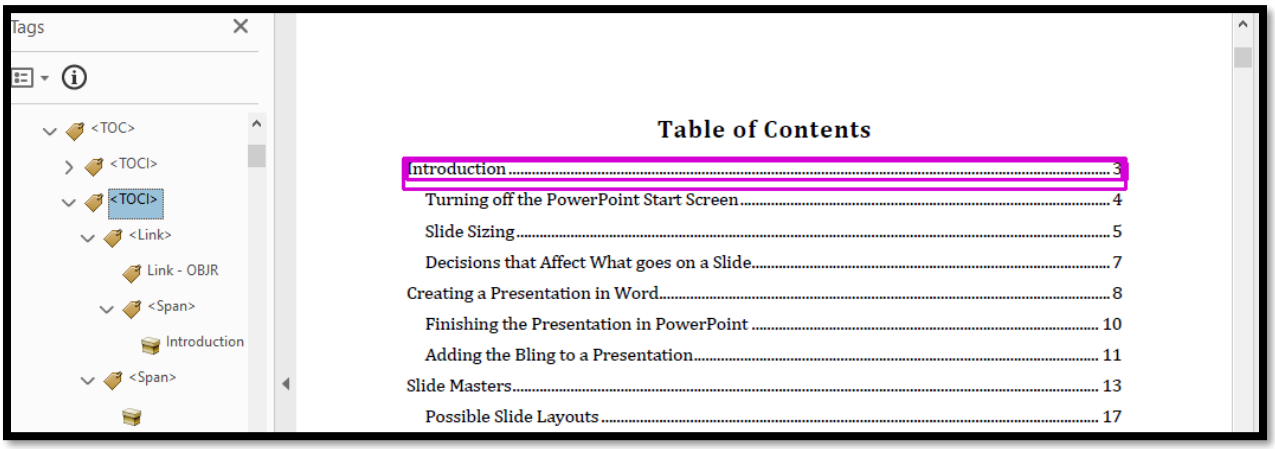

Figure 2 Table of Contents tagged from Microsoft Office using the Microsoft tagging tools.

Adobe Acrobat Pro DC and Microsoft Tagging Tools

What we have now, coming from both the Microsoft tagging tool and the adobe acrobat Pro DC conversion tool (Acrobat Ribbon in word) is a "dogs breakfast (scrappy mess) of inconsistent and unreadable Tags/Tables of Content.

For example, for a document I published in May 2019 which is the handout for my accessible PDF forms workshop, I got this mess when I got a list of the TOC using the JAWS screen reader.

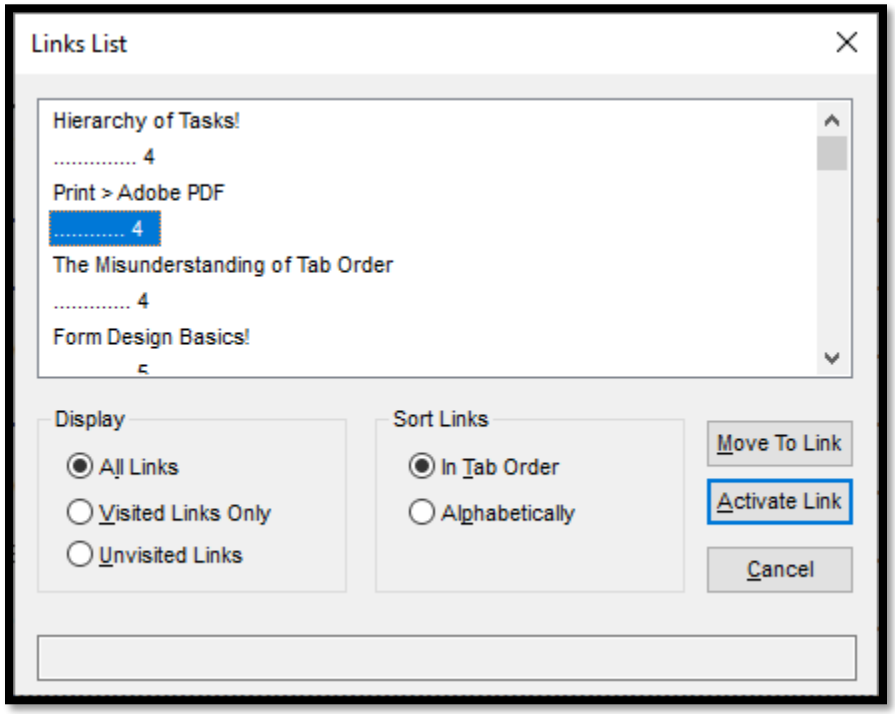

Figure 3 Unreadable Table of Contents in a PDF document from May 2019.

This is now "consistent" with what I get using an accessible Word document using either the Microsoft tagging tool or the Adobe acrobat Pro DC conversion tool from the Acrobat Ribbon.

The Tags Tree is even messier.

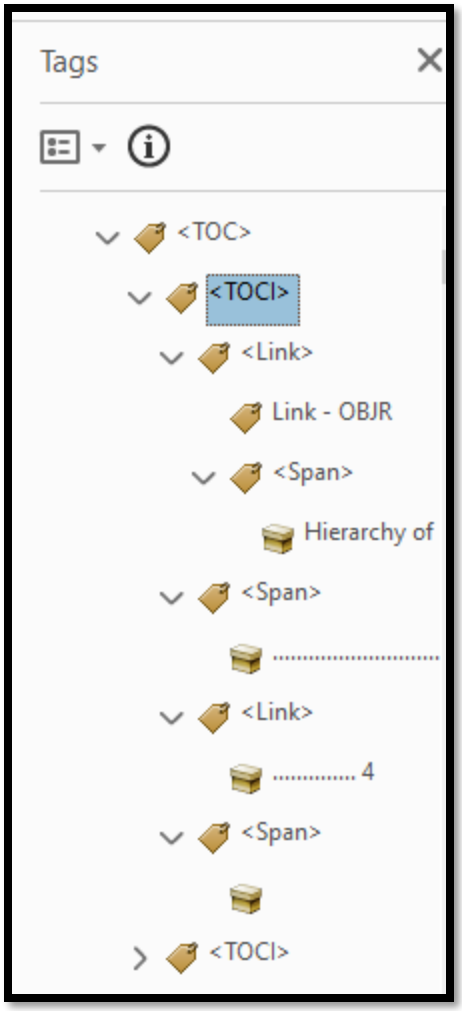

Figure 4 Tags Tree of a Table of Contents June 2019.

What does someone who uses a screen reader or Text-to-Speech hear? Instead of hearing a consolidated topic, dot, dot, dot, page number, we hear “Link, topic” “Link dot, dot, dot” “Link, page number.” None of this makes sense!

Those of us with disabilities should not have to piece together content so that it makes sense, Especially when the document was designed to be accessible in the first place!

If you have a visual disability and are using screen magnification, what are you supposed to do with the disconnected elements of a TOC?

If you have a learning, cognitive or print disability and your Text-to-Speech tool gives you a broken disconnected list like the one above, how are you going to use the TOC as a navigation tool?

This implementation creates a barrier to digital content where there wasn’t one to begin with.

So let's see what Foxit Phantom for Business and Nuance PowerPDF do with the same accessible Word document. There are files attached from both the Microsoft tagging tool and the Adobe Acrobat conversion tool as well as from Foxit and PowerPDF, but the results are the same no matter which tool you use.

Foxit Phantom for Business

I took the same accessible Word document and used the Foxit Ribbon in Word to do the conversion. The resulting PDF document opens in Foxit Phantom for Business.

Here is what we experience using the list of links with the JAWS screen reader in Foxit.

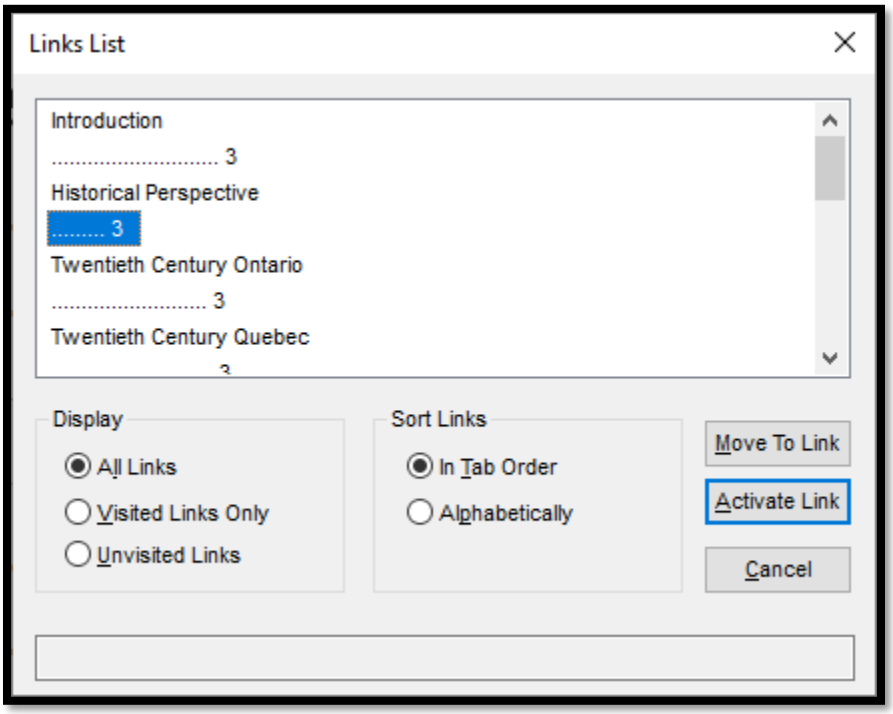

Figure 5 TOC list using the JAWS screen reader in Foxit.

Here is what the Tags Tree looks like:

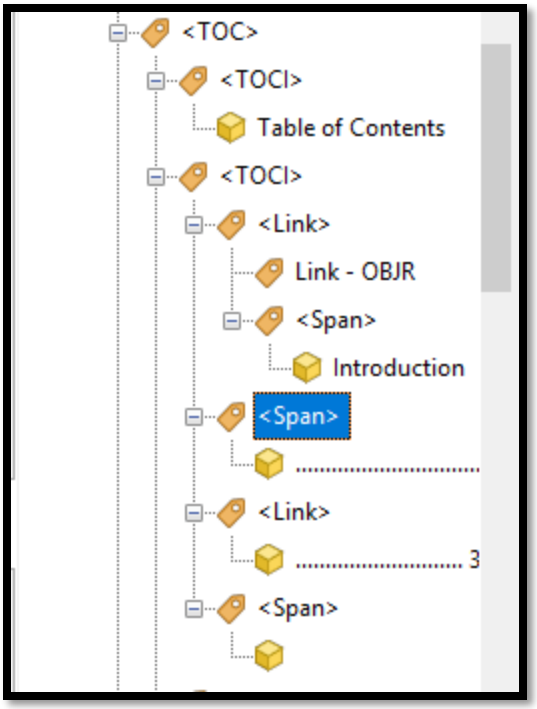

Figure 6 Table of Contents generated from accessible Word document using Foxit.

So we have the same mess of Tags when converting from Word using Foxit Phantom.

And I have to add the Tags Panel EVRY time I launch Foxit Phantom for Business.

Nuance PowerPDF

I used the Nuance PDF Ribbon in word to convert the accessible Word document to tagged PDF. The resulting PDF document opened in PowerPDF.

Using either the JAWS or NVDA screen readers, I can read the document but I can't get a list of links or a list of Headings because the PowerPDF keyboard commands over ride them. I can't use a JAWS or NVDA Pass-through key to bypass the assigned keyboard commands in PowerPDF . So, here is what we hear:

"Introduction, dot, dot, dot ,3" which is correct, followed by "Link, Link" which is not correct. "Historical Perspective, dot, dot, dot, 3" which is correct, followed by "Link, Link" which is not correct.

So the text and page number are not identified as a link and the links stand on their own with no connection to any text.

Since neither JAWS nor NVDA can give me a list of links in a PDF opened in PowerPDF, I switched to Adobe Acrobat Pro DC to get a list for a closer look.

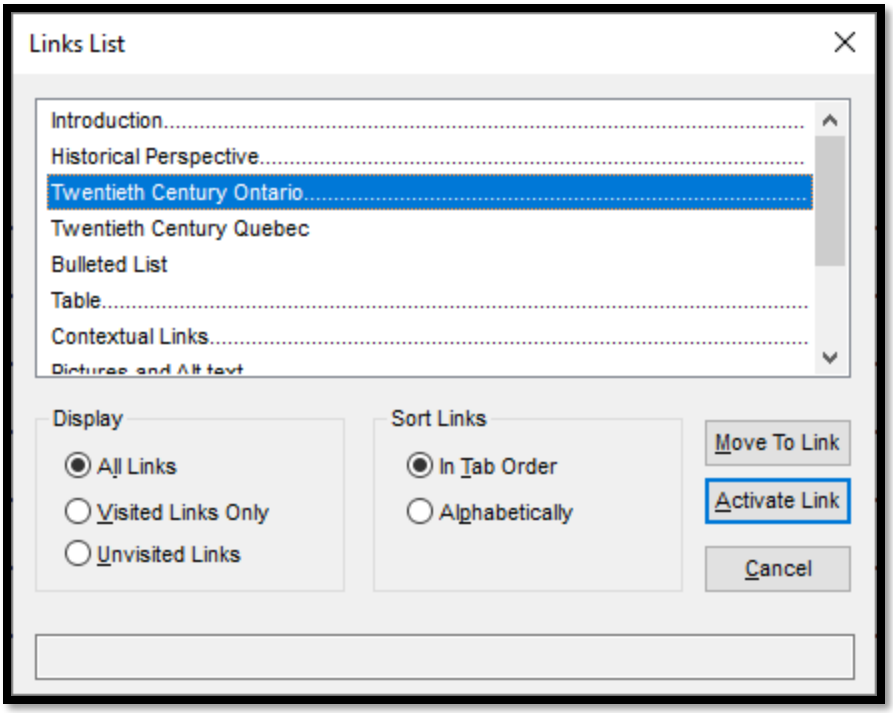

Figure 7 TOC list using the JAWS screen reader for PowerPDF document in Acrobat.

Here is what the Tags Tree looks like in PowerPDF.

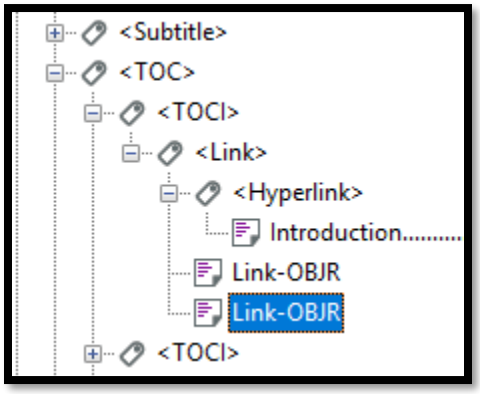

Figure 8 Tags Tree in PDF from PowerPDF Ribbon in Word.

Amazing! PowerPDF actually tagged the Subtitle correctly! It didn't put it in the <TOC>! I did use the Subtitle Style in Word for the text Table of Contents!

However, there are still problems with the tagging of a TOC. The <TOCI> Tag is still truncated.

Moreover, the Heading "Bulleted List" is not showing the Dot Leaders or the Page number when you get a list of Links using adaptive technology. If you look at the Tags Tree, one of the duplicate <Link> Tags is missing.

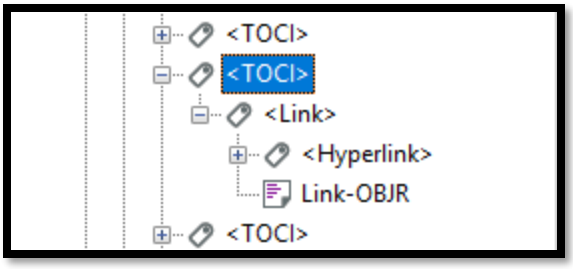

Figure 9 PowerPDF document showing "incomplete" <TOCI> Tag.

This didn't happen in the other test documents.

In PowerPDF you have to remember to turn on Highlight Content EVERY time you open a document and want to examine the Tags Tree.

While PowerPDF does read with a screen reader, you can't use your normal shortcuts like getting a list of Headings or a list of Links.

Summary

It is not clear what happened to the tagging of a Table of Contents from October 2019 to the beginning of 2019, but something dramatic happened with the implementation of PDF-1 and PDF/UA-1 that creates an accessibility barrier. Both PDF -1 and PDF/UA -1 are the current ISO specifications for PDF documents.

IF this is a preview of PDF-2, then either the specifications are not clear to developers or the specifications are missing something and are what is creating the accessibility barrier.

The fact that "everyone" is implementing the same broken tagging of a TOC and TOCI would indicate that this is a move toward some type of industry standard that has nothing to do with retaining and creating an accessible PDF document.

The Table of Contents in the sample Word document were accessible as of October 2018 and were read as a unified element (topic, Dot Leaders, page number).

Now, trying to navigate a TOC using adaptive technology is like trying to navigate a hot sticky mess on your carpet!

> **Note:** I did use NVDA to test the list of links for the TOC in each document and got the same results as I did with JAWS. PowerPDF does not let you use JawsKey + F7 to get a list of links or NVDA key + F7 to get a list of links, nor does it let you use any pass through tools to bypass the PowerPDF keyboard commands. F7 in PowerPDF opens the Spell Check dialog.

Appendix A: Contact Information

Karen McCall
info@karlencommunications.com

The Karen McCall School on Teachable[2] has for purchase items such as self-paced online courses and books on accessible Word, PowerPoint and PDF. The free tutorials are for Office 365 subscription (Office 2016 and 2019). You have to sign into the "school" but I only use the ability to send you e-mail to let you know that there is new content…no marketing.

The Karlen Communications website[3] has older tutorials pre Office 365 and older conference handouts as well as links to webinar archives.

All material, free or for purchase is in accessible tagged PDF. Videos in the self-paced courses are captioned.

[2] Karen McCall School on Teachable: https://karen-mccall.teachable.com/
[3] Karlen Communications website: https://www.karlencommunications.com/