

Table of Contents

Text in other languages	1
<i>Marking up Languages</i>	2
<i>For entire documents written in another language</i>	2
<i>Indigenous Languages</i>	2
<i>A note about poetry</i>	3
<i>Working with Images of Words and Different Alphabets</i>	3
Q&A Archive	3

Text in other languages

Whenever there is text in another language it's very important to properly identify the language of the text. This ensures that screen readers, braille displays, and other assistive technologies can render the content accurately and read the content according to the pronunciation rules for that language. When no other language has been specified for a phrase or passage of text, its human language is the default human language of the book.

In some cases, though, it's not desirable to markup the change in language as it actually negatively affects accessibility.

When there are frequent switches in languages in a book, the text-to-speech voice will also change, and this can be a bit jolting if it occurs frequently and depending on how different the voices are. For example, the reader might have "Apple Alex" set as the default English voice and "Apple Amelie" for the French voice. So, if it's not necessary to mark up the language, then it's often best to leave it. Just something to keep in the back of one's mind.

Do **not** mark up the language in these cases:

1. Proper names

1. Examples: Bellevue, Pierre

2. Technical and Scientific terms

1. Examples: Homo sapiens, Alpha Centauri, hertz, and habeas corpus
2. Most professions require frequent use of technical terms which may originate from a foreign language. Such terms are usually not translated to all languages. The universal nature of technical terms also facilitate communication between professionals.

3. Words or phrases that have become part of the language

1. Examples:
 1. "Rendezvous" is a French word that has been adopted in English, appears in English dictionaries, and is properly pronounced by English screen readers.
 2. "Podcast" used in a French sentence. Because "podcast" is part of the vernacular of the immediately surrounding text in the following excerpt, "À l'occasion de l'exposition "Energie éternelle. 1500 ans d'art indien", le Palais des Beaux-Arts de Bruxelles a lancé son premier podcast. Vous pouvez télécharger ce podcast au format M4A et MP3," no indication of language change is required.
2. Frequently, when the human language of text appears to be changing for a single word, that word has become part of the language of the surrounding text. Because this is so common in some languages, single words should be considered part of the language of the surrounding text unless it is clear that a change in language was intended. If there is doubt whether a change in language is intended, consider whether the word would be pronounced the same (except for accent or intonation) in the language of the immediately surrounding text.

4. Words of indeterminate language

1. In the rare case where, for one reason or another, we cannot determine what the appropriate language information is, then we just leave it as is (do not mark it up). This might be a situation where we're not sure if the text is non-linguistic. We haven't come across this situation in an ebook yet!

For more info please refer to the WCAG page on languages through this [link](#).

The important thing to keep in mind is why the guidelines exist. This guideline is for non-visual readers who use audio (text-to-speech) to access the text. I sometimes find it helpful to ask, “would this negatively affect reading comprehension if it were voiced in English or in French?”. You can easily test this out by activating the TTS on your Windows (Narrator) or Mac (VoiceOver)

Links for Windows Narrator:

- Video for Window Version 10: <https://www.youtube.com/watch?v=U25vhhE50kl>
- Written instructions for Windows: <https://support.microsoft.com/en-ca/help/22798/windows-10-complete-guide-to-narrator>

Links for Mac VoiceOver:

- Video for Mac: https://www.apple.com/voiceover/info/guide/_1124.html
- Video for Mac with Siri: <https://www.youtube.com/watch?v=HzMtcVS5c4E>
- Written Instructions for Mac: https://www.apple.com/voiceover/info/guide/_1124.html

Marking up Languages

To mark up secondary language:

- Select the text
- Go to Tools > Language
- This will open a pop up menu
- Select the appropriate language
- Apply Strong style to the word or phrase

When passing the ticket to the Production Coordinator, please make note of what languages you used.

The extra steps of applying Strong style and including a list of languages used in RT will help identify if they have been applied properly.

For entire documents written in another language

If the entire book is written in another language, we will need to change the language of the document so that it is not English.

To change the document language on a Mac, you can follow these steps: [Change document language on a Mac](#)

On a PC, Word should automatically detect the language of the document: [Change document language on a PC](#)

Indigenous Languages

Currently, we are not able to create a Language Style in Word for Indigenous Languages. In the future hope we will be able to create Language Styles, but for now we still want to be able to markup these

languages so they are set apart from the surrounding text.

There are two steps for marking Indigenous Languages:

1. Apply Emphasis style to the words and phrases the same way you would a Language Style.
2. Insert a Producer's Note at the beginning of the text to inform the reader what Indigenous Languages are in the book, and that Text-To-Speech is unable to pronounce these words.

It is important you try to include the proper names of the Indigenous Languages in the Producer's Note. Where you can, also include the Tribe name. Sometimes this is clear in the book, and other times you may need to do a bit of research. If you have any questions please contact the Project Coordinator.

A note about poetry

When you are working on poetry, you will **not** be able to apply a particular language style to words and phrases. In this case, you can just leave the Word version without language markup and use just the Poetry (Poem (DAISY)) style. Just make a note in the RT ticket that there are multiple languages.

Working with Images of Words and Different Alphabets

Sometimes a word or phrase will appear as an image in line with the sentence instead of typed text. This is a issue from the publisher. Words or phrases should not be formatted as images, but sometimes publishers do not follow these guidelines. When this happens you will need to transcribe the image of the term or phrase, and then apply the language style. Be sure to delete the images once you are done adding the text version.

Some languages cannot be transcribed due to the complexity of that language. An example would be Arabic. When it comes to languages like Arabic, unless you are a native speaker you cannot transcribe it correctly. In this case you would treat the image of the word like other images in the document and add Alt-Text stating it is an Arabic Word. You would then put a Producers Note at the beginning of the book to explain why you did this. If you are unsure if the language is something you can safely transcribe please contact you supervisor for more feedback.

Sometimes the terms or phrases are typed out in line with the rest of the text, but with a language that uses a different alphabet. In this case, if the text appears as typed text, and not an image, then you can simply apply a language style to it as usual.

In case you're not sure how to type in different languages, this is how you do it on a Mac [Enable keyboard layouts in different languages in Office for Mac](#) and [Windows](#).

In other cases you can use unicode to enter the characters of the language. For more information on unicode go to the [Symbols](#) page.

Q&A Archive

Q: I'm working on the play "1 Hour Photo." It contains a few Japanese characters but in the

conversion, the characters were changed to Roman alphabet letters instead. The English translation is given for the symbols so I'm wondering if I should just erase the Roman alphabet letters. Or would it be better to insert the proper ideogram back in? If so, how do I do that?

[Here is an example: Tetsuro raises both hands to illustrate the ideogram for "mountain," III.]

Another option I thought of was to copy the image of the ideogram from the PDF file and paste it into the Word file. Then, add alt-text to it. What do you think?

A: You should insert the proper ideogram back in. You can do this using unicode. Here are [the instructions on how to set that up](#)—but remember, some [languages are too complex for this technique](#). If you feel confident you can insert the correct ideogram, then do so. Remember, we **never** have text as images, even if it is in another alphabet.

Q: That's the thing, I don't know how to find the correct Japanese ideogram in Unicode. I don't even know which Japanese alphabet to search in - apparently there are several. I don't feel at all confident that I can identify the correct symbol. I know how to insert symbols with Unicode - the missing part is how to identify the specific code for the correct Japanese symbol. I think it would be one of the CJK Unified Ideographs but I don't know which one and I can't just search "mountain" to find the correct one. The instructions you point to on the wiki don't explain that part. To me, this falls under "Some languages cannot be transcribed due to the complexity of that language" which is why I was wondering if I should find a work-around to still include the symbols for people who do understand Japanese. Or, just leaving the symbols out since the English translation as well as the English pronunciation of the Japanese word are both included.

A: In this case, since it is an issue of conversion and you are not confident in finding the correct ideogram, then simply put a producer's note at the beginning of the book explaining that the original Japanese ideograms did not convert to this version of the text, but the translation and punctuation are present—or something better written than that to explain the issue.

Q: I am editing an illustrated children's book that has a sentence where I think I need to indicate a foreign language. It is just a single word but it is clear that a change in language is intended (Page 3 of *The Gathering* by Theresa Meuse). I tried to follow the instructions for creating a new style but the Mi'kmaw language is not one of the language options. What should I do?

A: Unfortunately, there are currently no language tags for that language. What you can do is put a Producer's Note in the book with something like "This book includes words and phrases in Mi'kmaw language. Text-to-speech software will not be able to pronounce these words and phrases correctly in the Word version."

Q: I have a book that uses Innuinaktun words, but it also has two images. One is an image of a table with the word symbols beside the sound (no English translation), and the other is a full piece of text in Innuinaktun. How should I address these images in the Alt-Text? And should I also include a producer's note about the Innuinaktun words?

Symbol	Sound	Symbol	Sound	Symbol	Sound	Finals	
Δ	i	▷	u	◁	a		
∧	pi	>	pu	<	pa	<	p
∩	ti	⋈	tu	⊂	ta	⊂	t
ρ	ki	∂	ku	⊂	ka	⊂	k
∩	gi	∩	gu	⊂	ga	⊂	g
Γ	mi	∩	mu	⊂	ma	⊂	m
σ	ni	∂	nu	∂	na	∂	n
∩	si	∩	su	∩	sa	∩	s
⊂	li	⋈	lu	⊂	la	⊂	l
∩	ji	<	ju	∩	ja	∩	j
∂	vi	∩	vu	∩	va	∩	v
∩	ri	∩	ru	∩	ra	∩	r
∩ρ	qi	∩∂	qu	∩⊂	qa	∩⊂	q
∩∩	ngi	∩∩	ngu	∩∩	nga	∩∩	ng
∩	lhi	∩	lhu	∩	lha	∩	lh

∩∩∩∩∩∩∩∩∩∩

∩∩∩∩∩∩∩∩∩∩

∩∩∩∩∩∩∩∩∩∩

∩∩∩∩∩∩∩∩∩∩

∩∩∩∩∩∩∩∩∩∩

∩∩∩∩∩∩∩∩∩∩

∩∩∩∩∩∩∩∩∩∩

∩∩∩∩∩∩∩∩∩∩

∩∩∩∩∩∩∩∩∩∩



A: Looks like this is the Inuktitut language, according to the publication information. Inuktitut can be represented by [Unicode Canadian Aboriginal Syllabics](#). We will need to translate the images into Unicode. If you're using Mac, enable your "Unicode Hex Input" keyboard (see Language section in wiki for instructions). To type each symbol/letter into Word, hold down the alt key and type the 4-digit



number, i.e. 1400.

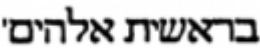
Q: I am editing a poetry book that uses Italian, French, and Latin. If I apply a language to one word, it changes the entire line or stanza. Should I just leave it as poetry style?

A: Unfortunately, identifying languages in Word doesn't translate well to DAISY XML and requires manual editing of language tags in the XML. You can just leave the Word version without language markup and use just the poetry style. Just make a note in the RT ticket that there are multiple languages.

Q: I have a book that deals with hebrew words. Some of the words are typed, and I can create a style for them, but the other words are images of just a letter, or an entire word. How should I deal with them? Should I just put in the alt-text this is an image of this hebrew letter/word? Or should I put in a producers note? I included examples below:

Greek is  pronounced  I
int with pebbles, to compute, calculate,
/ reasoning * The same word is found in

There are over a dozen specific Greek arti
only one, and that is the letter  It means the
becomes part of the word that it modifies. The
Hebrew that acts similar to the article and is al
the particle . It appears in the Old Testame

Gen 1:1 

A: Using images instead of text is a very bad publishing practice :(Images of text should all be converted to text in the body of the narrative. We should type out all the text including the Hebrew and Greek text and use a style to tag them as words in the Hebrew or Greek language (as we usually do with foreign language words).

In case you're not sure how to type in different languages, this is how you do it on a Mac [Enable keyboard layouts in different languages in Office for Mac](#) and [Windows](#).

WCAG 2.0 - H58: [Using language attributes to identify changes in the human language](#)

From:
<https://bcl.wiki.libraries.coop/> - **BC Libraries Coop wiki**

Permanent link:
<https://bcl.wiki.libraries.coop/doku.php?id=public:nels:etext:language>

Last update: **2020/06/30 09:36**

