# Transcription Style Guide

## About this document:

This style guide has been developed by SpokenWeb team members during the production of *The SpokenWeb Podcast*. It reflects our formatting guidelines, as well as process tips, workflow, and evidence of collective decisions. It is a living document that we continue to add to as a team based on conversations, research, and new transcription examples that come up in podcast episodes. We also aim to gather resources on transcription best practices and include them here.

Do you have any transcription questions or suggestions for improvement? We'd love to hear from you and include contributions from our community to strengthen our transcription practices. Email us at spokenwebpodcast@gmail.com or reach out on social media @SpokenWebCanada.

Style guide last updated: May 2022

Contributors to date:
- Natasha Tarr
- Stacey Copeland
- Hannah McGregor
- Judith Burr
- Kelly Cubbon
- Katherine McLeod

## Guiding principles:
- Transcriptions can take multiple forms depending on the core aim of the transcription (accessibility, citation, SEO, etc.)
- The primary aim of SpokenWeb transcriptions is accessibility.
  - We do this by releasing completed transcripts simultaneously with audio episode releases.
- Transcription work is ongoing and incomplete. There is always room for improvement and we can always learn from best practices others have established. However, it's prudent to use some standardization to streamline and implement our current understanding of transcription practices.

# Process tips:

- Transcribe in Temi, using the highlights function for placeholders for questions or key things to adjust (i.e titles to italicize later, spelling checks for names)
  - The notes section (located on the left hand side in Temi) can be a good place to jot down and time-stamp notes as you go, to then summarize in an email to the supervising producer when delivering the draft transcript.
- While listening and transcribing, re-read transcriptions to view them from the mindset of a reader, rather than a listener.
- Export from Temi into Word doc with Columns
- Arrange columns so the order is time stamp, speaker name, content.
- Remove time stamp hyperlinks (see shortcut [here](#))
- Copy Word doc to Google doc
- Update speaker names in full (Temi has a character limit that often cuts off audio clip descriptions)
- Use this style guide to format the Google doc transcript in full.
- Use Google doc comments to highlight any outstanding issues, leave questions, ask for clarification etc. so the episode producer(s) can respond directly.
- Notify the Supervising Producer/Project Manager that the transcription draft is up in the Google drive folder and ready for review.
- Be on standby for edits, and respond to them.

# Formatting

<u>Speaker Names</u>
The speaker should be labeled with their full name. Be sure to spell check by googling or asking.

- For recorded audio, performances, and archival audio, we typically aim for the following formatting.
  - [Name, Location/Event, Year]

- Sometimes the same person will speak in different contexts, such as in a voice over, an interview, and an archival recording. Be sure to differentiate their voice in the speaker names so it is clear for a reader of the transcript.
  - I.e  Kaie Kellough, Zoom interview, 2022
    AND Kaie Kellough, Words and Music Festival, 2019

Body
The document should be left justified, with the names of each speaker aligned to the left margin, and the transcription text beginning 1.5" from the left margin. This happens naturally when you download with the output format as "Columns".

There should be a double space after the podcast intro and a double space between the text of each new speaker. Include a paragraph break (double space) when the host changes topic or segues to the podcast outro.

For accessibility purposes, aim to break up big blocks of text. If one person is speaking or performing for a long time, you can create a line break and repeat the speaker name with an updated timestamp. This is easier on the eyes for reading and makes it easier to reference specific timestamps.

Mistakes
If someone makes a mistake when quoting someone or saying the name of a book, show, etc, add [sic: should read as "WORD"] in square brackets.

For example, "And we read the play by Linda Griffiths *Alien Creature: A Visitation by* [sic: should read "*from*"] *Gwendolyn MacEwen*."

# Grammatical Functions

- Exclude transcription of filler words like "um" and "er," but include "like," "so," "sort of." Exclude "you know" when it is said with a similar cadence as other filler words such as "like" and "so".
  - This formatting decision is based on creating a clearer, easier to read transcript.
  - It may be worth revisiting this in the future, particularly as we balance honoring people's speech patterns and tendencies.
    - An interesting resource to consult: [Why Captioning Non-Speech Elements Matters for Accuracy](#)

- Exclude the extra word(s) in speech repetitions. For example, in "I've, I've learned along the way," delete one one of the instances of "I've".
  - Include repetitions with *variation*. For example, "I will say, I'll begin by saying" and "It was like, it's like Sappho, it's like Sappho wrote a novel."
  - Include repetitions when used for *emphasis*. For example, "Many, many tapes have come our way."
- Use the Oxford comma because it's great, cool, and fantastic.

## Hyphens
- Include a hyphen after negative prefixes if the word isn't already in the common lexicon without a hyphen.
  - ex: anti-capitalist, non-heteronormative
- Include a hyphen after the prefix "re" if the word isn't already in the common lexicon without a hyphen OR if the root word begins with "e."
  - ex: rewatch, re-educate, re-politicize

## Capitalization
- When turning a proper noun into a verb (ex: I googled it, I tweet at her all the time) don't capitalize the verb.
- Capitalize major and minors for studies. For example, Computer Science.

## Acronyms

- If something is introduced and then later referred to by an acronym, include the acronym in square brackets beside the first full mention of the term.
  - ex: UBC Okanagan [UBCO]

## Titles
- *Italicize* titles of Books, Podcasts, Movies, TV Shows, Albums, Video Games, Board Games, Art Shows, Art Pieces, etc.
- Put "quotation marks" around titles of Articles, Chapters, Episodes, Songs, Poems, Podcast Episodes.

## Numbers
- Centuries: 20th century
- Years: '70s, 1970s, 80th birthday
- Age: "I was in my twenties"
- Zero to nine spelled out, 10 and up written numerically.
- Dates: January 23rd, June 9th
- Time: 5:00 AM, 7:00 PM

## Quotations
- When the speaker quotes someone else or themselves, use typical quotation style.
  - ex: I was like, "Whoa, a live reading of a poem is totally different than the reading of a poem on a page."

- When quoting poetry, separate each line with a "/" with spaces on either side
  - ex: "His pampered Muse / Knew no veto. / Hers lived / When his Muse cried / He replied / Loud and clear / Yes! Yes! I'm waiting here."
  - *NOTE: if a speaker is quoting their own poetry, no quotation marks are needed.*
  - *NOTE: if there's no written copy of the poem, do not try to guess where the line breaks (/) go. Simply write out the poem with punctuation that sounds right to you.*
  - *NOTE: Please check in with the episode producer to see if there are any copyright concerns about replicating written versions of published works.*
  - *Tip: Try Googling the poem to see if there is a printed version in existence online. If there isn't, proceed with punctuation that sounds right to you unless the episode producer communicates otherwise.*

Pauses and Interruptions
- For short and long pauses, use a comma.
- Use an em dash (—) with a space on either side to separate a speaker's parenthetical thoughts from the primary sentence content.
  - ex: When I came to Concordia — it's almost like in this sort of fortuitous moment — I was asked to have a meeting with the department chair in the first week that I was installed in the department, just to go over very practical things.
- Use double hyphens where a speaker is cut off by another speaker, and another set of double hyphens to indicate where the first speaker picks back up.
  - ex: Jason Camlot:     And then one morning, my professor, John Miller--
    John Miller:        I'm John Miller, and I've retired from teaching Victorian literature at Concordia University in Montreal.
    Jason Camlot:     --walked in with a boombox.
  - When multiple people speak over each other, format as follows, naming specific people if they are known and using the tag "Various Voices" if not.
    - Various Voices:     [Overlapping] Hello. Hi. Hi.
  - When an unknown person speaks, label as "Unknown Person [number]" with number differentiating one unknown voice from another. For example:
    - Unknown Person 1:    Can you hear me?

# Inclusive terminology
- Include a space between the prefix "trans" and the rest of the identity description
  - ex: trans man/male, trans woman/female, trans masc/femme, etc
- Include a space between the prefix "cis" and the rest of the identity description

- ○ ex: cis man/male, cis woman/female
- There should be no space following the prefix when transcribing the general concepts relating to gender/orientation.
  - ○ ex: cisgender, heteronormative, transphobia
- Capitalize Black and Indigenous when referring to people, communities, cultures, etc.
- For names pertaining to First Nations, adhere to that nation's accepted transliteration.

## Audio Clips from Other Media

- If the audio clip is *distinctly separated* from the speaker, include a paragraph break (double space) and format it in the two columns Temi provides, with the title of the audio in the "Speaker" column and the description in the transcription text column.
  - ○ ex: Audio Recording:   [inaudible] My boy Ikie is a bright boy. Some day he shall set the moon on fire. [inaudible]

- If the audio clip is *interrupting* the speaker in the middle of a paragraph or sentence, include it in square brackets and transcribe any audible words.
  - ○ ex: So I think, for example, about an author reading in 1970 and a now defunct gallery of a hub mall on the University of Alberta campus. [Audio, from the 1970 recording: Please come in and sit down, there's lots of room at the front.]
  - ○ *NOTE: if the audio clip is not clearly introduced, describe it briefly in the square brackets, as shown above.*

- Note: The nature of this podcast means there are often episodes with overlapping audio clips, archival audio of performances or readings, sound poems, experimental performances with non-verbal sounds etc. These can be challenging and time consuming to transcribe. Remember the transcription is something that will be read, so the transcription should indicate what is occurring but doesn't need to reflect verbatim beat by beat what is happening. Use concise descriptions to aid understanding where possible.
  - ○ Ex. [Begin: Sound Poem Performance. Overlapping Non-Verbal Sounds Interspersed with Repeated Snippets of Poetry]

# Non-verbal audio

Music

There is often instrumental music throughout the podcast.

- If the music is *distinctly separated* from the speaker, include a paragraph break (double space) and format it in the two columns Temi provides, with the title of the audio in the "Speaker" column and the description in the transcription text column.
  - ex: Music:    [Instrumental]
  - *NOTE: where possible, describe the kind of instrumental music that is playing. For example,* [Instrumental Piano]
  - From S3 onwards, episode producers will provide their scripts which should have music cues with specific song names or sound samples. We've asked them to do this to expedite and standardize the process of transcribing music cues.

- If the music is in the *background* or is *interrupting* the speaker in the middle of a paragraph, include it in square brackets and describe as usual, marking the places where it begins and ends as follows:
  - Ex: Just like that story that Jason described earlier of finding a box of tapes and eventually listening to them. [Begin Music: Upbeat Instrumental] …. But for now, let's go back to Michael who shared with us the process of working with this kind of material. [End Music: Upbeat Instrumental]

- When music changes volume or is 'played under' other audio for lengths of time it can be challenging to decide how to note these changes. With the aim of accessibility and legibility, please include significant musical changes with descriptions such as the following:
  - I.e [Music Continues Under Audio] or [Music Continues as Musical Bridge]


Theme music

It should be formatted on its own line with the two columns Temi provides, with the title of the audio in the "Speaker" column and the description in the transcription text column.

The standard way to format full episode theme music is:
- *SpokenWeb Podcast* Theme Music:  [Instrumental Overlapped With Feminine Voice] Can you hear me? I don't know how much projection to do here.

The standard way to format ShortCuts theme music is:
-  [Piano Overlaid With Distorted Beat]

Sound Effect

- Format as [Sound Effect: Description]
  - ex: [Sound Effect: Box Opening]

Action

- Any specific and recognizable sound in the original recording (ie. not a sound effect put in during editing) should be described.
  - ex: [Whistles], [Beep], [Chiming Bells]

Laughter

- This style guide originally excluded instinances of laughter, but we have since reversed this decision. Laughter is communication and provides context for the tone of an exchange. It can also be a valuable indication of how audiences are responding to performances in recorded and archival audio.
- Include laughter
  - When a speaker laughs note it using square brackets
    - [Laughs]
  - When an audience laughs in response to something note it like this:
    - [Audience Laughs]

# Commonly Used Terms and Phrases

- *SpokenWeb Podcast*
- #spokenwebpod
- @SpokenWebCanada
- What does literature sound like? What stories will we hear if we listen to the archive? Welcome to the *SpokenWeb Podcast*: stories about how literature sounds.
- SpokenWeb's Audio of The Month: 'mini' stories about how literature sounds.
- Audio of the Week/Month
- minisode, meta-podcast, mini-podcast
- ShortCuts
- SPOKENWEBLOG

TIP: Review the episode summary prior to starting the transcription in Temi to see key names, titles, and references.

- Ex. https://spokenweb.ca/podcast/episodes/alone-together/