General Transcription Guidelines

Casing

All words and names, without exception, must be transcribed in lowercase at all times.

Examples:

- **i** need directions to **montreal**
- **maria** goes to **walmart** in **seattle**
- do you know **google chrome**

Only acronyms and spelled-out letters should be capitalized. When a letter is spelled- out, it must be capitalized, followed by a period, then by a space.

Examples:

- i'll be available from nine **A. M.** until five **P. M.**
- all my data is stored in a **D. V. D.** and several **C. D. s**
- las diez imágenes en formato **J. P. G.** are stored in my **U. S. B.**

Acronyms that were created as initialisms, and are conventionally written in all caps, but are pronounced like regular words, should be transcribed like regular words

Examples:

- the U. S. A. has been part of **nato** since nineteen forty nine
- he is without a doubt the best **fifa** player

Numbers

All words numbers must always be spelled out.

Examples:

- if necessary call **nine one one**
- i was born in january seventh of **nineteen eighty four**
- these ruins are over **one thousand and five hundred** years old

Special Characters

Special characters, sometimes used for short-hand notation, should not be used in transcription.

The following are examples of **incorrect** transcriptions:

- coffee & tea
- i'll arrive @ five P. M.
- you are the #1

The following are examples of **correct** transcriptions:

- coffee **and** tea
- i'll arrive **at** five P. M.
- you are the **number** one

Named Entities

Names of people, companies, songs, countries, apps, etc. should all be transcribed in accordance with previous rules (e.g., use lower case). Copyrighted terms (e.g., trademarks, brand names, registered names) are also to be treated the same way.

Well-known people, companies, songs, apps, etc. should all be transcribed with the trademarks or official spellings that they are known by.

If you are unsure how to accurately spell one of these named entities, then perform a web search to identify the correct spelling.

Examples:

- OneDrive → **onedrive**
- McDonald's → **mcdonald's**
- PowerPoint → **powerpoint**
- Spider-Man → **spider-man**
- Coca-Cola → **coca-cola**
- Inglourious Basterds → **inglorious basterds**
- O Canada → **O. canada**
- Ice-T → **ice T.**

If the correct spelling is still not clear, then make your best judgement based on the context that you have.

Math Symbols

Math symbols must never be used. Instead, the Word associated with the symbol must be used.

Examples:

- 2 + 2 = 4 → two plus two equals four
- 6 / 3 = 2 → six divided by three equals two

Informal words

If a speaker clearly uses an informal pronunciation that is **Formally Recognized** as a word then use the informal word.

Examples:

- you **wanna** do that instead
- are they **gonna** come at the usual time
- I **kinda wanna** tell you something

But do not create new words.

Examples:

- **Full of → fulla**
- **Like a → lika**

If the speaker explicitly says "**ha ha ha**" (which is different from legitimate laughing, which the guidelines consider as noise), "bow", or "bang", etc. transcribe them if, and only if, the dictionary considers it is a **formally recognized** word.

Ungrammatical Words

For words which are ungrammatical in the given context, but are clearly articulated, transcribe them as spoken.

The following are examples of **correct** transcriptions:

- there you will find all my **document**
- **where** was the last time you called her
- **is** you completely sure about it

Please note the sentences above are grammatically incorrect given the context. Yet, if these words were spoken this way, these would still be considered **correct** transcriptions.

Ambiguity

Ambiguity can come in many forms:

- **Homophones** (words with same sound but different spelling)

Examples:

- hear / here
- bear / bare
- bee / be

- **Variant Spellings** (words with more than one acceptable spelling)

Examples:

- ax / axe
- doughnut / donut
- barbecue / barbeque

- **Proper Nouns**

Examples:

- Vanessa / Vanesa
- Anthony / Antony
- John / Jon

For homophones, variant spellings and inflections, simply choose the form that you feel is most likely given the context.

For proper nouns, you may use the most commonly used variation.

Commonly Used Foreign Words

If the foreign word/language is part of the language's regular lexicon (i.e., would be understood by most speakers), write it in the foreign language using foreign language script.

The English word **Okay** is one of most used borrowed word in other non-English languages. The transcription of it should be **O. K.** when it is used as borrowed word in other languages.

Active Tags

<UNKNOWN/>

This tag is to be used in the following scenarios:

- **Unintelligible Speech**

Use this tag when there is obvious human speech, but one or more of the actual words cannot be determined. **Unintelligible speech** may be a result of strong accent, speech is too fast, poor pronunciation, poor sound quality or any other reason.

For cases where part of the speech is intelligible and part the speech is unintelligible, use **<UNKNOWN/>** for the unintelligible portion and continue to transcribe the intelligible portions as usual, including tags.

Note: Do not use consecutive (uninterrupted) <UNKNOWN/> tags. If the speaker says several unintelligible words in a row, only one <UNKNOWN/> tag is required.

- **Foreign Speech**

  - Partial-Utterance Foreign Speech

If part of the utterance is in a foreign language, then mark that part of the utterance as **<UNKNOWN/>** but transcribe the remaining words as usual.

  - Full-Utterance Foreign words

If the entire utterance is in a foreign language, use the **<UNKNOWN/>** tag to represent the entire utterance. But if the speaker mentions a foreign **named entity**, transcribe it.

- **Truncated Speech**

Use the **<UNKNOWN/>** tag when either or both the first word and/or the last word in an utterance is cutoff (**truncated**) by the start or the ending of the recording.

Note: For cases where the entire utterance is **<UNKNOWN/>** (either unintelligible or in a foreign language), don't add any additional tags. Simply tag the entire utterance with a single **<UNKNOWN/>** tag.

Note: Do not use this tag when the speech is clear but ambiguous due to various reasons, such as homophones, variant spelling, etc.

- **Mispronunciations**

A **Mispronounced Word** is defined as a word the speaker uttered incorrectly, yet its true intent can be deduced by the context of the utterance.

For mispronunciations where the intent is clear, the standard word should be used.

- los angeles is in **californa** (doesn't pronounce the 'i') → los angeles is in **california**
- open **misocroft** office (due to the context, we can infer the intent) → open **microsoft** office

For mispronunciations where the intent is <u>not</u> clear, the **<UNKNOWN/>** tag should be used.

<FILL/>

A **<FILL/>** tag must be applied for both words which are not fully spoken and uttered sounds which lack any semantic value.

Incomplete words

Use the **<FILL/>** tag when a word is damaged and not fully uttered by the speaker.

Examples:

- I would like a burger and a **sod**… actually just a burger → I would like a burger and a **<FILL/>** actually just a burger
- please call me tomo… tomorrow → please call me **<FILL/>** tomorrow

- **Filler**

Use **<FILL/>** tag if the speech includes sounds without semantic value, such as **umm**, **amm**, **ehh**, etc.

Examples:

- **amm** what would you like for dinner → **<FILL/>** what would you like for dinner
- **umm** perhaps we should **umm** change our plans → **<FILL/>** perhaps we should **<FILL/>** change our plans

- **Exclamations and interjections**

- **Repetitions**

Repetitions must not be tagged if the words are not damaged; simply transcribe the words as uttered by the speaker, regardless of the number of repetitions.

Examples:

- i i do not know what to to to say → **i i** do not know what **to to to** say
- are are are you sure about that → **are are are** you sure about that

Note: If the speaker uses official words as a filler (such as "like", "I mean"), do transcribe those words.

<CNOISE/>

If a **Continuous Noise** is present through the majority of the recording, at a volume around or above the speech's volume level, **<CNOISE/>** tag must be placed at the **beginning** of the transcription.

Examples:

- Continuous ambient noise (rain, traffic, etc.)
- Continuous electronic noise
- Continuous background music
- Continuous mechanical sounds

- **Singing**

For transcription purposes, **singing** will not function as actual speech. Instead, singing is to be considered a form of **Continuous Noise**. And so, it must be excluded from the transcription and instead represented by the **<CNOISE/>** tag.

If an utterance includes both speech and **singing**, the speech must be transcribed as usual while the **singing** will be represented by the **<CNOISE/>** tag.

Note: If a recording includes only **singing** and/or music, simply tag the entire utterance with a single **<CNOISE/>** tag.

Note: If a **Continuous Noise** is present in the recording, but stays well below the speech's volume level through the majority of the recording, it must not be tagged.

<SN/>

A **Sudden Noise** is defined as a short duration noise which occurs at a volume around or above the speech's volume level and must be represented with the **<SN/>** tag, according to its position in time, and as many times as it appears during the recording.

Examples:

- Short duration human noise (cough, sneeze, loud breath, etc.)
- Short duration electronic noise (loud crackling, popping, etc.)
- Short duration mechanical sounds (doorbell, car horn, etc.)
- Short duration sound effects

As mentioned above, the **<SN/>** tag may be placed along the transcription as described by the following scenarios:

- **Isolated Sudden Noise**

If a sudden **isolated sudden noise** occurs independently of uttered words, then it should be tagged independently (**with spaces**) of the transcribed words:

Examples:

- he's about to ring the bell **<SN/>** did you hear it

- **<SN/>** there are too many **<SN/>**

- **Simultaneous Sudden Noise**

If the sudden noise overlaps with a word, append the tag (with **no spaces**) to the end of the word.

Examples:

- I will definitively**<SN/>** call you tomorrow

Note: If a **Sudden Noise** is present in the recording, but occurs well below the speech's volume level, it must not be tagged.

Note: **<CNOISE/>** and **<SN/>** tags are not exclusive from each other and both may be applied in the same transcription as long as both of them are distinctively present and required.

<BA/>

This rarely used tag must be applied as the only tag in a transcription when a recording presents completely **Blank Audio**, absent from any type of sound signal.

Transcription of Multiple Speakers and Speech Overlap Tag

Some recordings include two or more speakers interacting with each other or the recording device. In these scenarios, all speech must be transcribed within the same line, regardless of the number of speakers, and according to their position in time.

Example:

- **Speaker I** : **word1 word2 word3**   **word4 word5**
- **Speaker II** :   **word1 word2**
- **Speaker III**:   **word1 word2 word3**

Expected transcription:

word1 word2 word3 word1 word2 word4 word5 word1 word2 word3

Background Speech

Background Speech is defined as speech which was not intended as part of the recording but is still audible. All forms of speech, including **Background Speech**, must be transcribed according to the rules above.

Example:

- **Speaker I** :   word1 word2
- **Speaker II** :   word1 word2 word3
- **Background Speaker**: word1 word2   word3 word4

Expected transcription:

word1 word2 word1 word2 word1 word2 word3 word3 word4

<OVERLAP><OVERLAP/>

Use the **Overlap** tag to encapsulate a section of the recording which presents two or more speakers at the same time.

Unlike other tags, this pair of tags must go around the overlapped speech section as shown below.

Independent Speech **<OVERLAP>Overlapped Speech<OVERLAP/>** Independent Speech

The louder, predominant, most notable speech will be transcribed, while the speech which goes behind the main speech will be ignored and not transcribed.

If no speaker appears to be louder or more important than the other, simply use your best judgment based on the context.

Example:

Overlap

- **Speaker I : word1 word2 word3 word4 word5**
- **Speaker II**:   **word1 word2** word3 word4 word5

Expected transcription:

word1 word2 word3 <OVERLAP/>word4 word5<OVERLAP/> word3 word4 word5

UNKNOWN, **FILL** and **SN** tags are still applicable inside the **<OVERLAP><OVERLAP/>** capsule. So do insert them whenever necessary.

Example:

Overlap

- **Speaker I :   word1 word2 word3 word4 word5 word6 word7 word8 word9**
- **Speaker II: word1 word2 <FILL/> word3 <UNKNOWN/>**

Expected transcription:

word1 word2 word3 <OVERLAP/>word1 word2 <FILL/> word3 <UNKNOWN/><OVERLAP/>