(a) Model

Target Text Embeddings

T_1 T_2 T_3 T_4

LLM ❄️

S_1 S_2 S_3 S_4 | P_1 P_2 P_3 P_4 | T_1 T_2 T_3

Speech Embeddings   **Prompt Text Embeddings**

CTC Loss

Projector 🔥   CTC Output

Pre-trained WavLM 🔥   Fine-tuned WavLM ❄️   LLM Tokenizer

Speech: Speech: Prompt: Transcribe speech to text. Some hotwords might help. The hotwords are {}.

Inference time   *Training time*

Stage 1   *Stage 2*

(b) Hotwords Generation for Training

Input

Shall I **compare thee** to a summer's day 🎲
Thou **art more lovely** and more temperate
Rough winds do shake the darling buds of **May**
And summer's lease hath all too short a date

Hotwords

Compare, thee, art, more, lovely, May

(c) Hotwords Generation for Inference

Input

Ground Truth: hello **bertie** any good in your mind
CTC Inference Sentence: hello **bertye** any good in your mind
Bias List: [achiet, anty's, ..., **bertie**, bords, ..., **hello**, ...]

1️⃣ **Build Index Dict**

'ac': {'micaceous', 'vacates', 'achiet'}
'ch': {'mutch', 'achiet'}
'hi': {'pethick', 'achiet'}
...

2️⃣ **Remove Common Words**

hello bertye ~~any good in your mind~~

hello bertye

3️⃣ **Retrieve Candidates**

{'itineraries', 'sojers', 'savonnerie', 'khem', 'inattentively', 'fortalices', 'humility', 'longville', 'plop', 'hello', 'ravelston', 'bryerley', 'letthers', ...}

4️⃣ **Match Similar Hotwords**

hello bertye
⬇️   ⬇️
hello bertie

Hotwords

hello, bertie