No-fluent Speech 🎤   🖥️ **Instruction:** What do you think of the pronunciation?   **Reference Text** Please Call Stella   Accented Fluent Speech 🎤

| GPT 4o (Real-Time) | Your pronunciation of "please call stella" was quite clear. The main thing is to focus on "l" sound in "please" and "Stella". | Your pronunciation of "please call stella" was quite clear. |

| SSDM | For the pronunciation of "please call stella", there is a stutter on the "c" sound in the word "call" from 1.52 to 1.60s and a prolongation of the "s" sound from 2.00 to 2.10s. | Your pronunciation of "please call stella" was quite clear. |

No Reference Text Required

| Ours — SSDM 2.0 | For the pronunciation of "please call stella", there is a stutter on the "c" sound in the word "call" at 1.48s. There is a stutter of "s" sound from 2.10 to 2.30s. | For the pronunciation of "please call stella", there is a phonetic error on the phoneme "IY" in the word "please" from 1.28 to 1.32s. |

Higher Accuracy, 10X training complexity reduction than SSDM, In-context "Pronunciation" Learning