| Rhythm | Pitch | Content | Timbre |

→ 1st round encoding and reconstruction for the original utterances.

→ Random factor substitution and speech reconstruction of the substituted factors.

⇢ 2nd round encoding for the speech recovered from the substituted factors.