

① aesthetics → character  
joy

VI

attach  
emotion  
visualization  
?

Input text

Finish/return

If words  
finish add  
extra expressions

break up  
words

expression  
determines  
speech style

Mouth open

- smiley (pitch: 1~2, rate: 0.5~1) ▼
- frowny (pitch: 0.5~1, rate: 0~0.5) ▲
- neutral (pitch: 1, rate: 0.5) -

Open mouth  
each time  
pronounce  
a word

Continuously

Eye Blink? tongue out?  
"HaHa"  
one eye blink  
left  
right

Left/right ○

- eyeBlinkLeft m
- eyeSquintLeft -
- eyeWideLeft ○



jawOpen "HaHa"

- mouthSmileLeft ☺ ▼
- mouthFrownLeft ☹ ▲



- browDownLeft ~ ^ ▲
- browInnerUp ^ ^ ▼
- browUpperUpLeft ^ ^ ▼

- cheekPuff ○ ○ ▼
- cheekSquintLeft ○ ○ ▼

tongueOut U "lol"

V2

if tongueOut, insert

else jawOpen >



update  
caption?

changes  
color of  
a button?

○ → rate?

Realtime? Simulate realtime ...  
language?

Send button  
triggers speech  
breaks up text  
based on punctuation

between part,  
capture facial expression,  
initiate new speech if  
text is not done

insert lol  
haha  
present

P: 1-2  
r: 0.5  
P: 0.5~1  
r: (0-0.5)