Write Out Loud

A multi-modal way to learn Chinese characters and more

Junru Ren and Freya Tan
Final Presentation

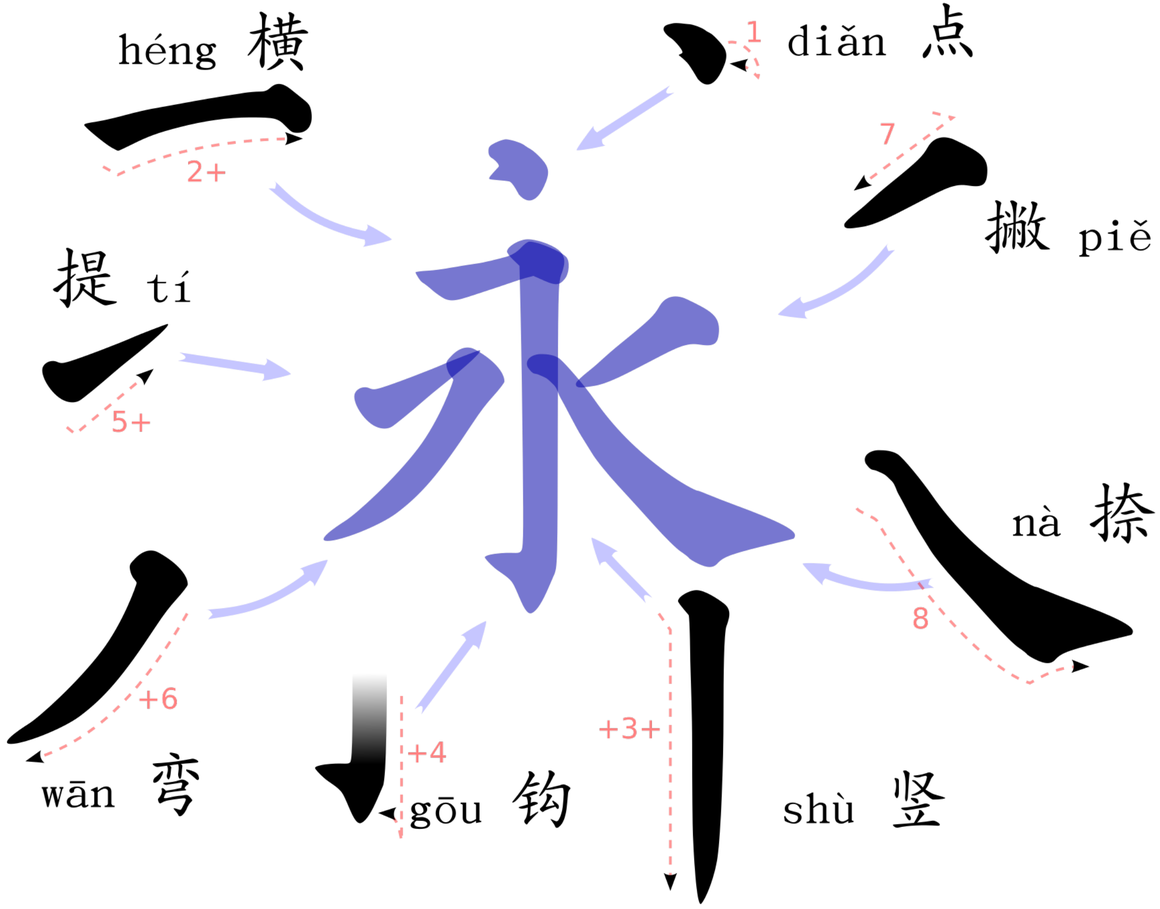
Traditionally...

Writing Chinese characters is learnt in silence

Yet...

Learning a new character is multi-modal

- Its pronunciation(s)
  - How does it **sound**
  - How to romanize it (pinyin, etc)
- Its shape, decomposed into
  - Different strokes
  - In a **particular** stroke **order**



Source: https://commons.wikimedia.org/wiki/Commons:Stroke_Order_Project/Kangxi_radicals#/media/File:8_strokes_of_%E6%B0%B8-zh.svg

Guide a Chinese learner to
vocalize each stroke
while **writing** a character

“Multisensory Learning”

Demo


11:37 PM Wed May 7
WhtsOutLoud

Show Real-time Transcript

Characters Selection

口 人 日

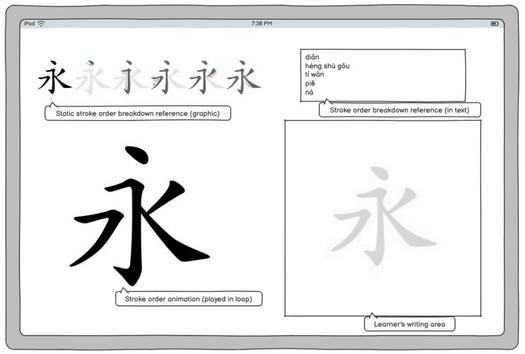
kǒu - mouth

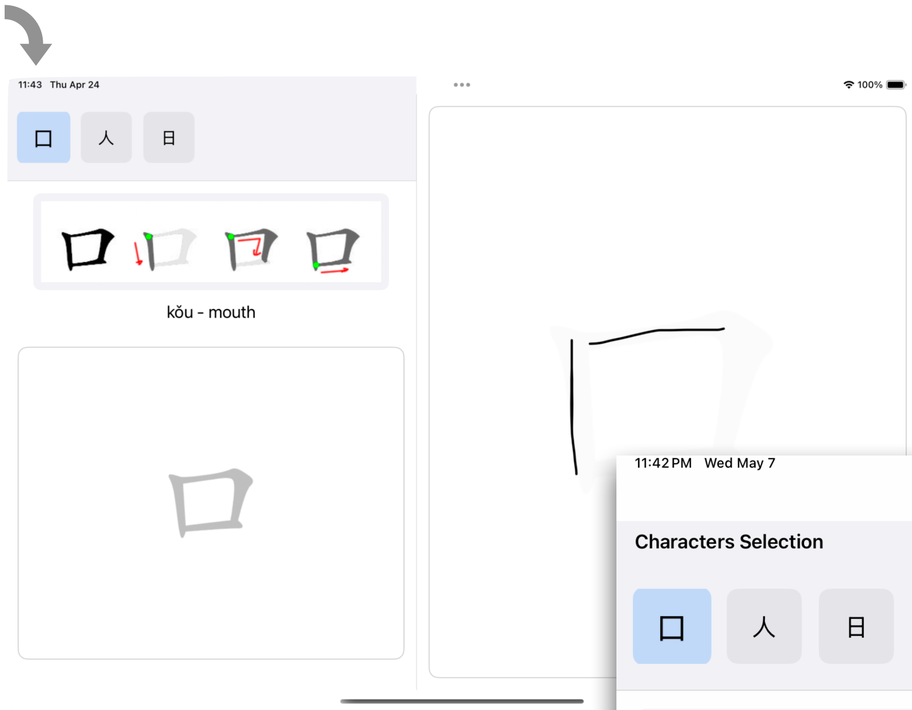
Vocalize your stroke while writing:

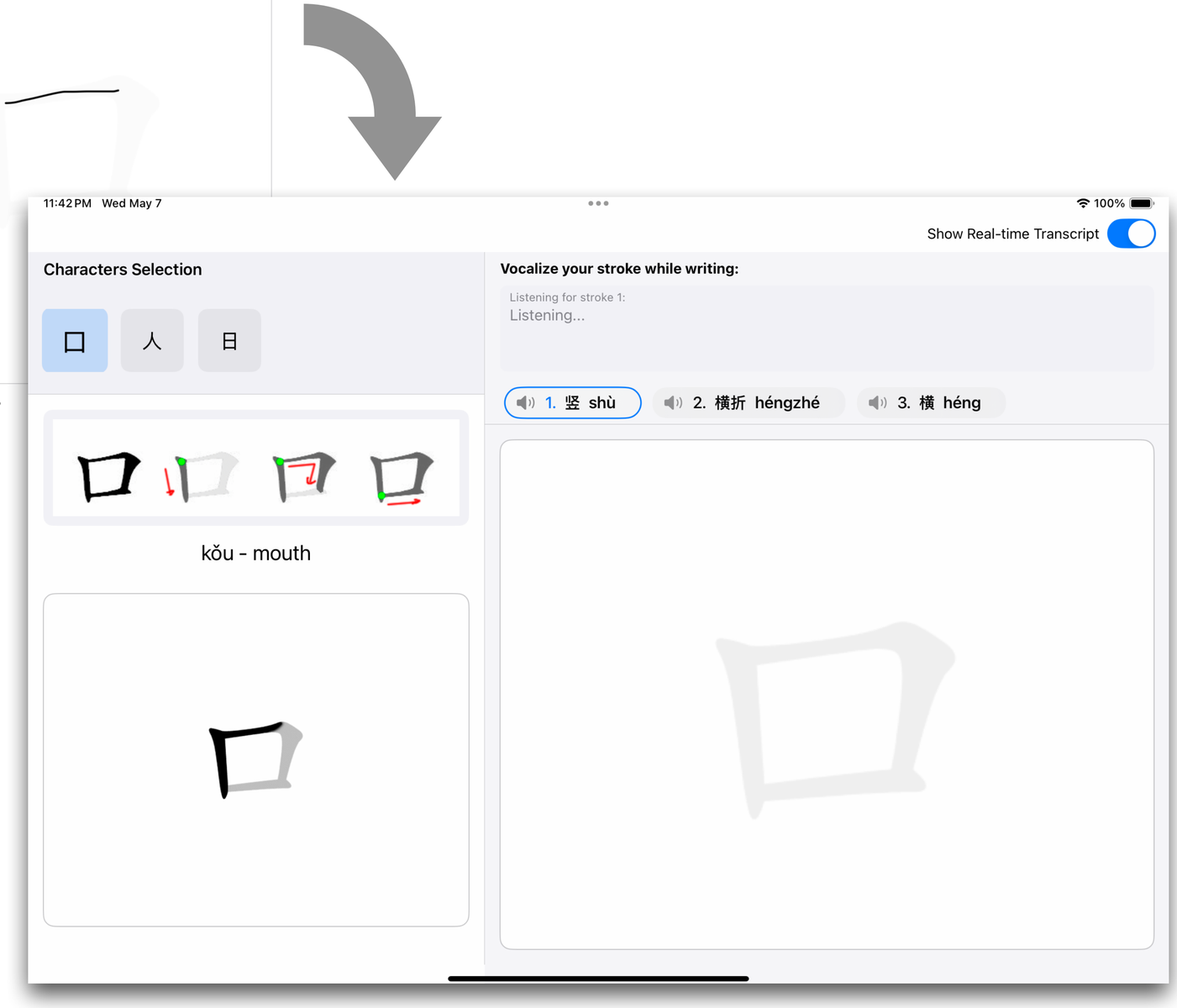
Listening for stroke 1:
Listening...

1. 竖 shù | 2. 横折 héngzhé | 3. 横 héng

Updates

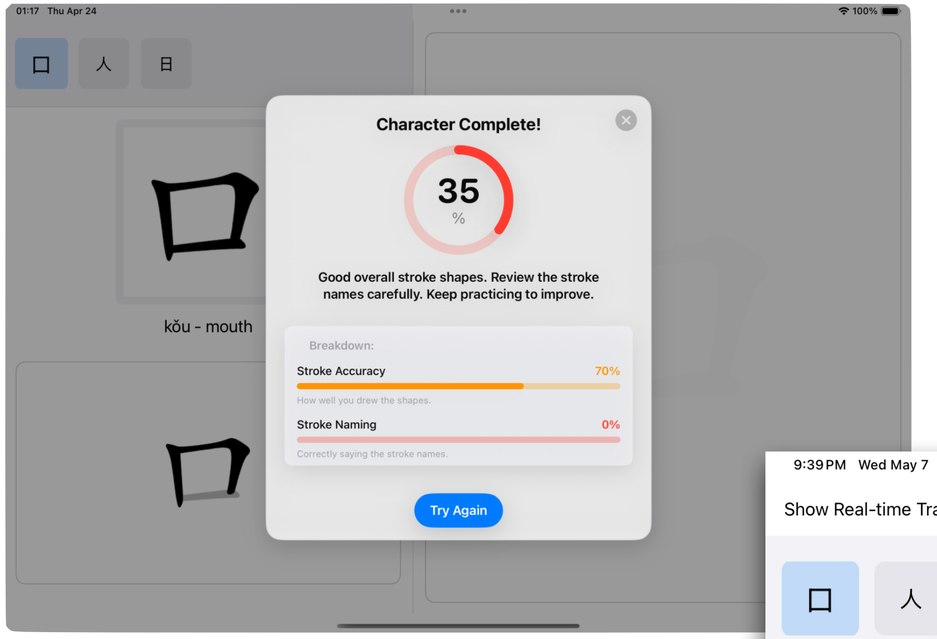


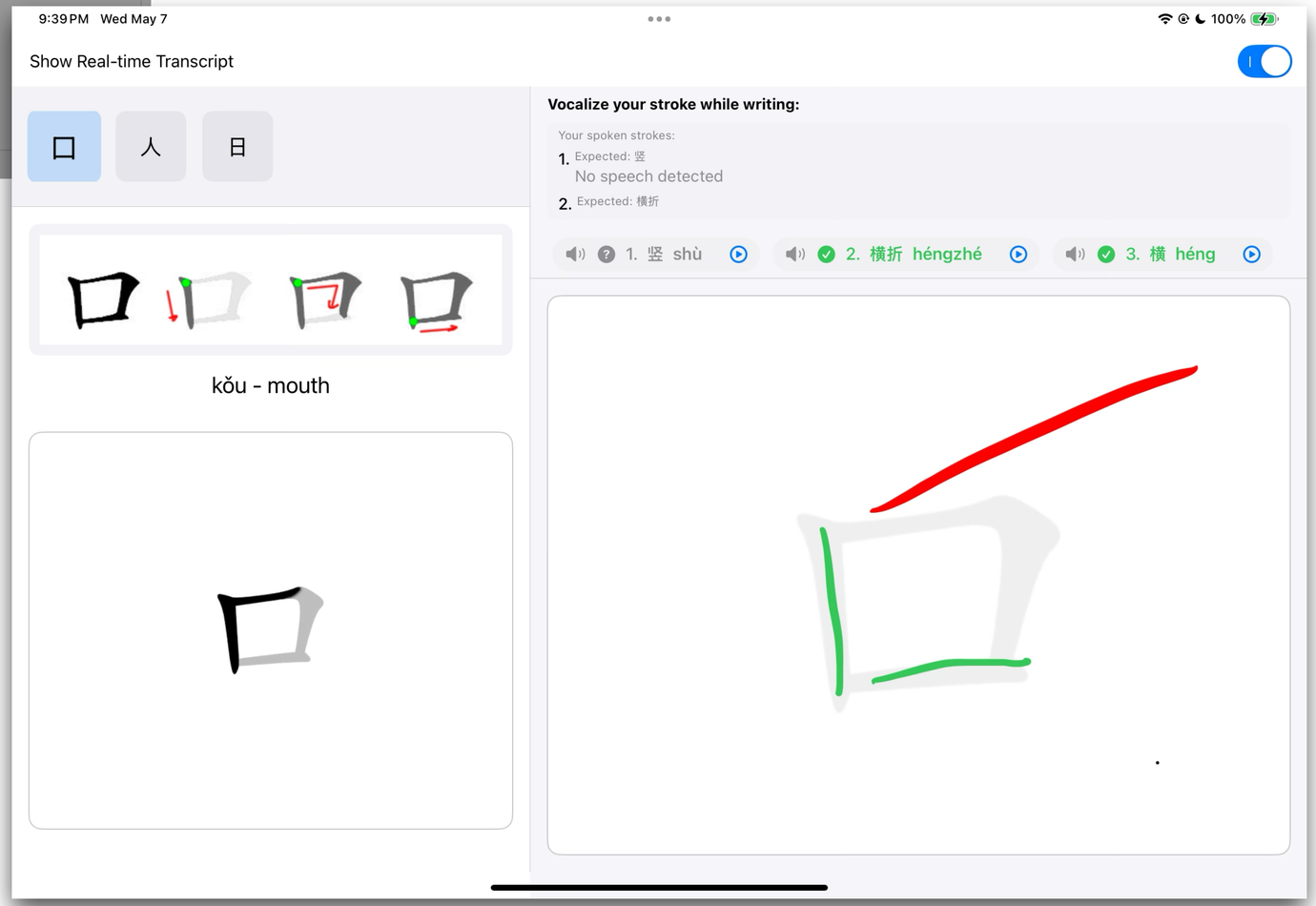




Enriched Layout

Just enough info to get started



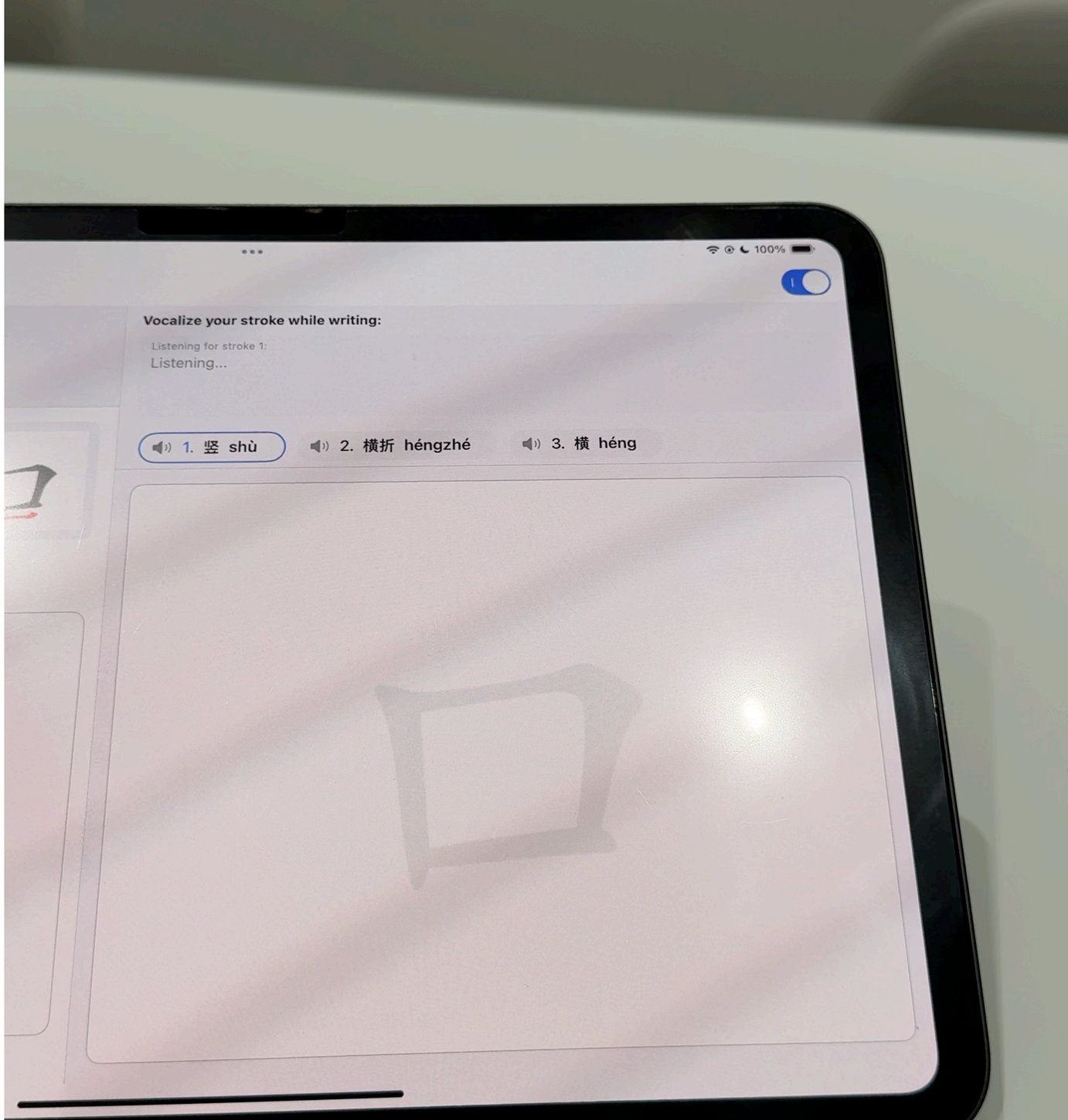


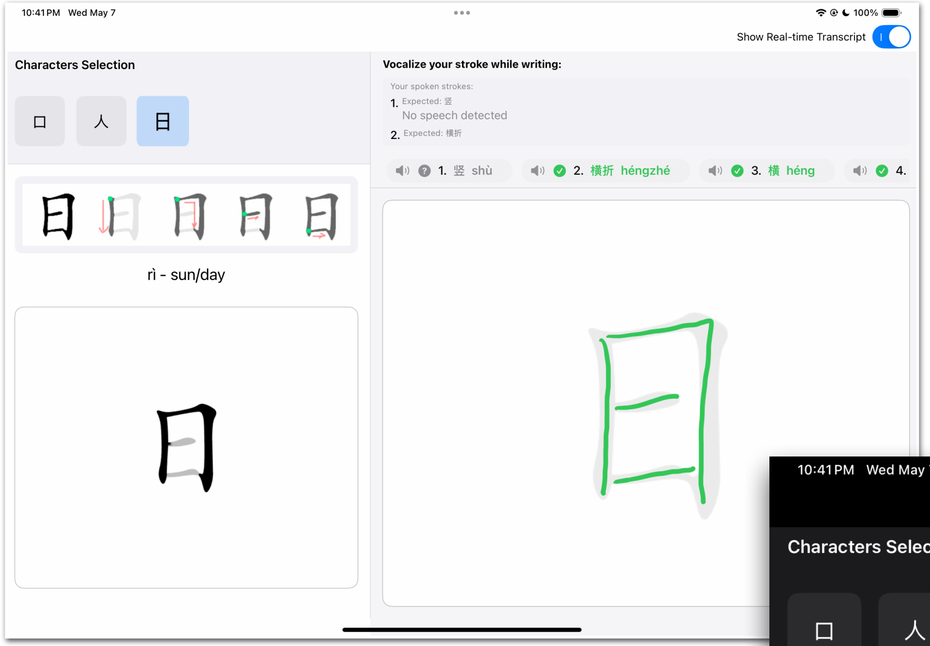
Simple Feedback

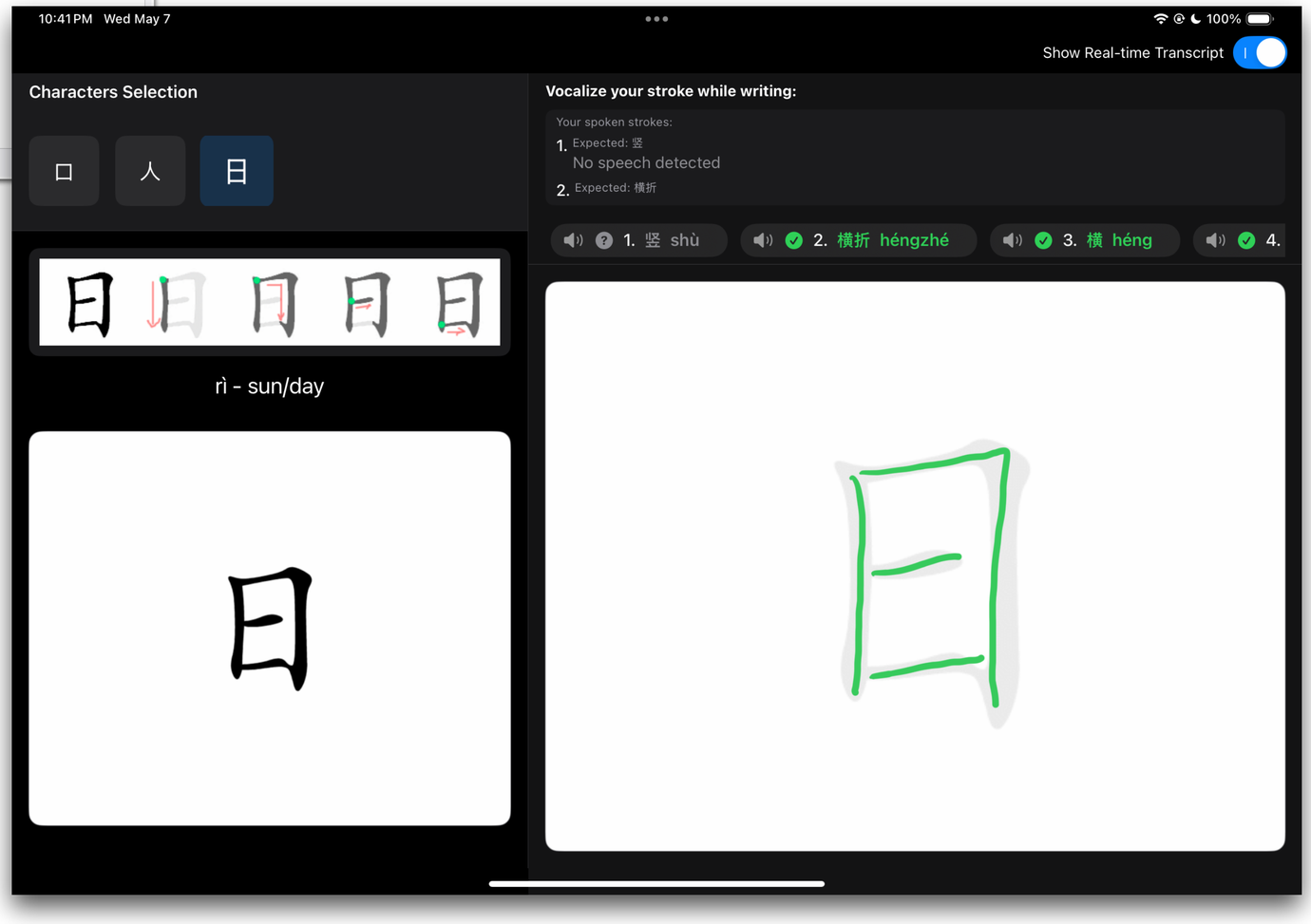
Correctness visualized on the spot

New:
Stroke Name Example

Made possible by
Apple’s native (robotic) Text-to-Speech 🤖







Dark Mode

Because why not...

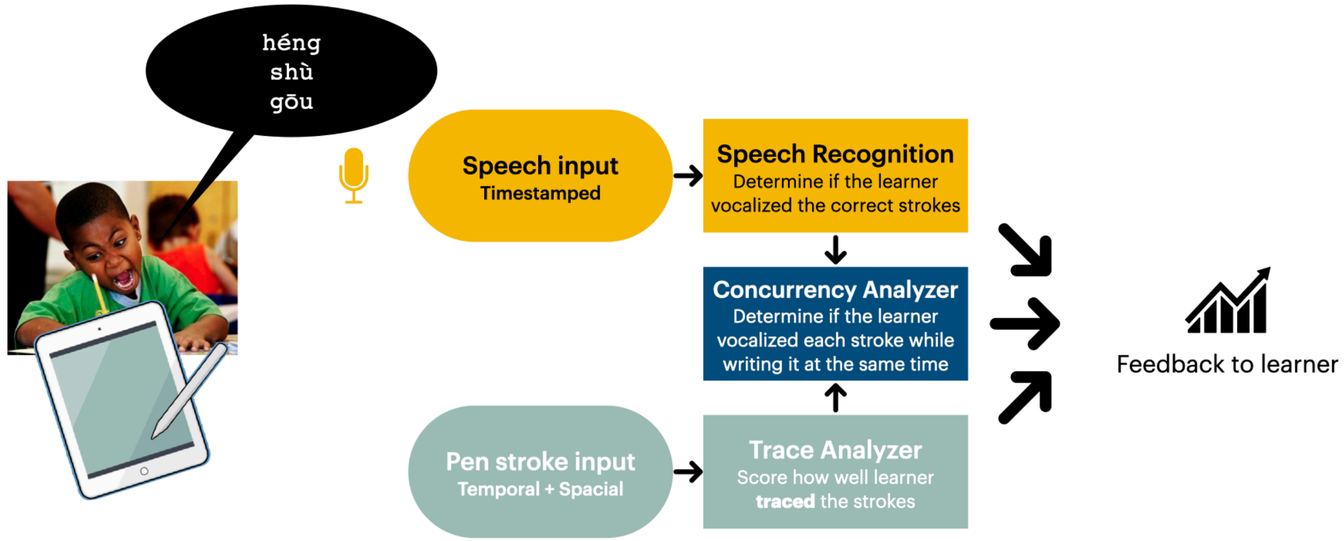
How it works

Apple SDKs & a lot of heuristics

- **PencileKit** for pen traces
- **SFSpeechRecognizer** for speech recognition
- **AVFoundation** for example pronunciations
- **WebKit** to render gifs

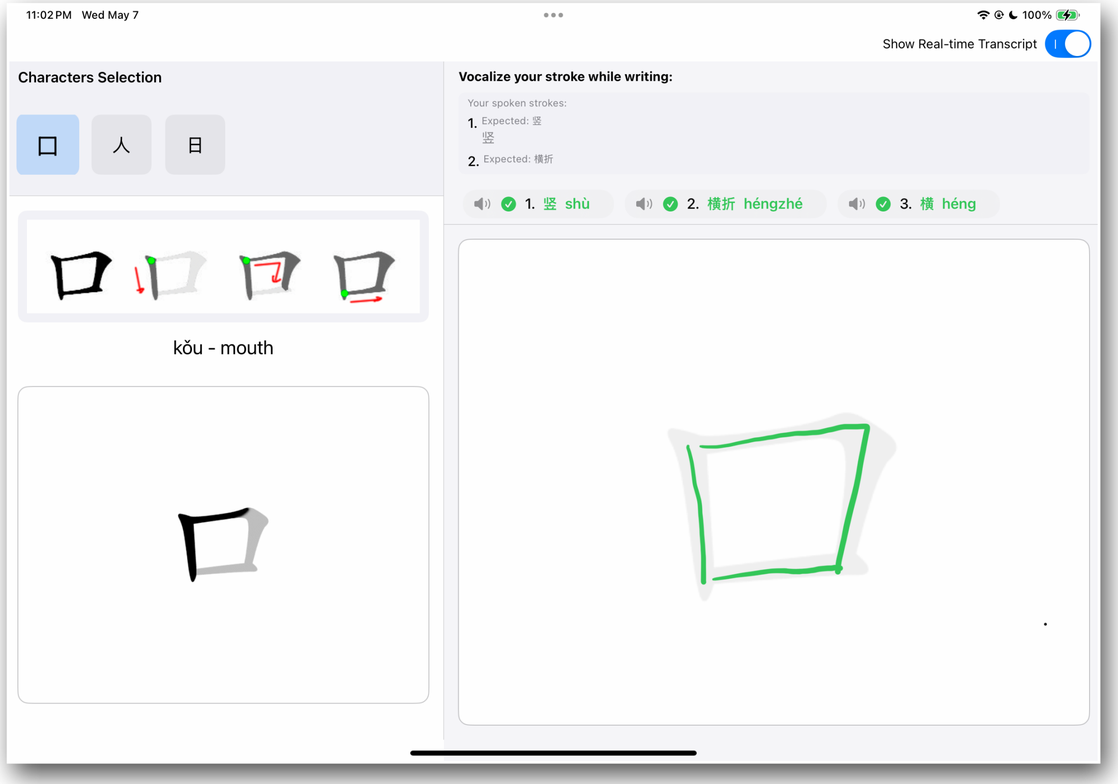
System Diagram

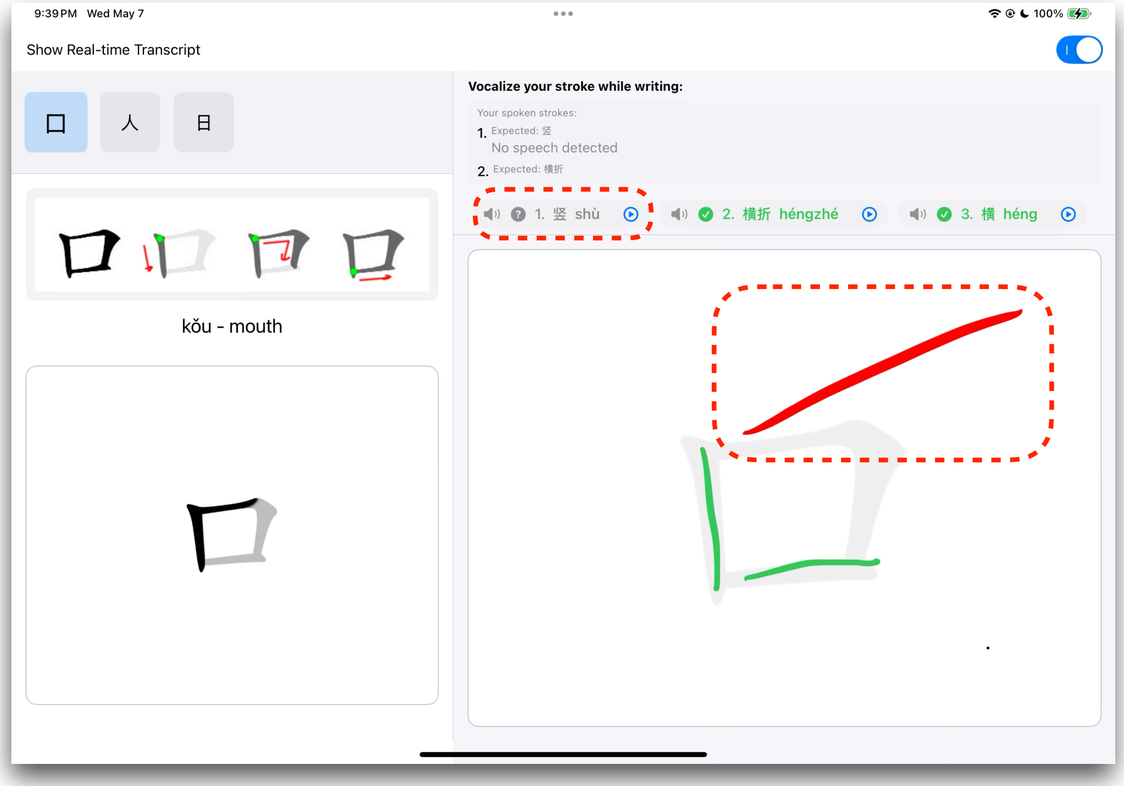
Learner traces through a displayed stroke and vocalize it



How well it works

Bad strokes and vocalizations can be caught



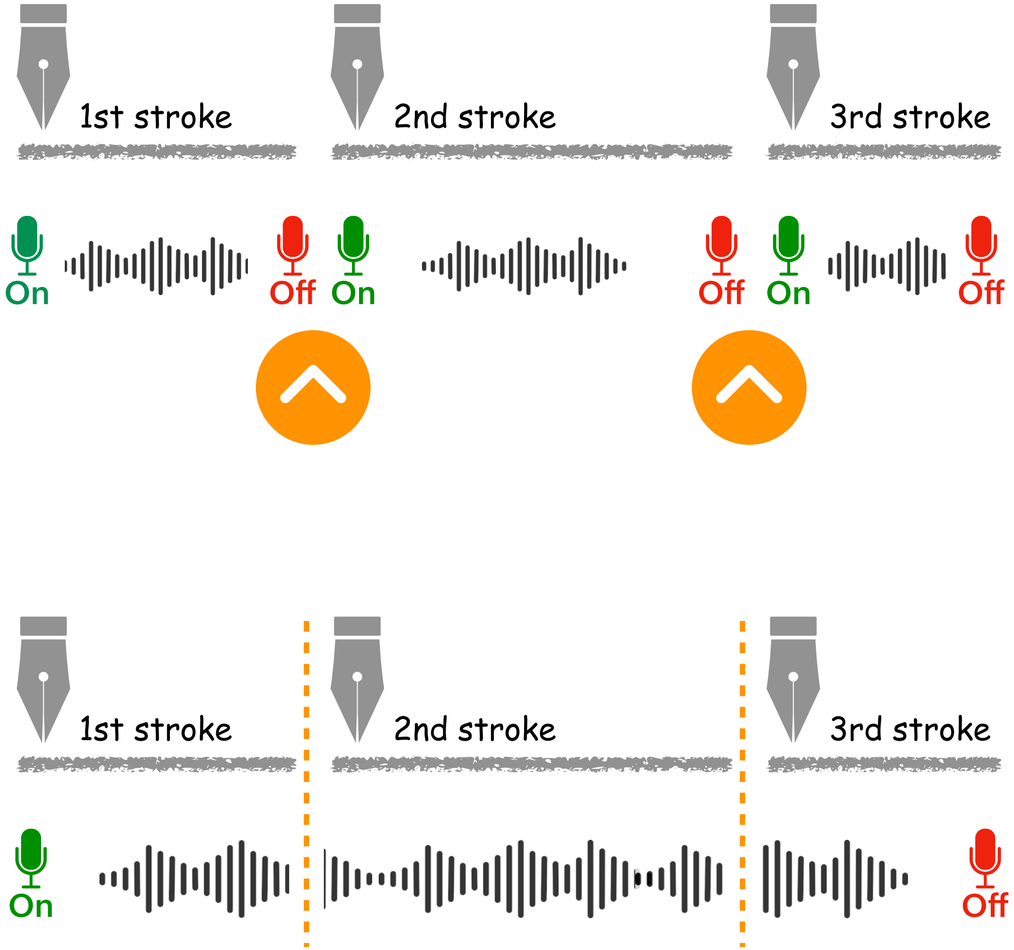


Learnings

Interesting Failure

Timing of speech segmentation

- Originally, speech recognition was started and stopped with **each stroke**
- Mic toggling is **a point of failures** → terrible speech recognition on each audio clip
- Now, speech audio is collected **continuously**, and then segmented **retroactively** for uttered stroke names
- Result: improved reliability



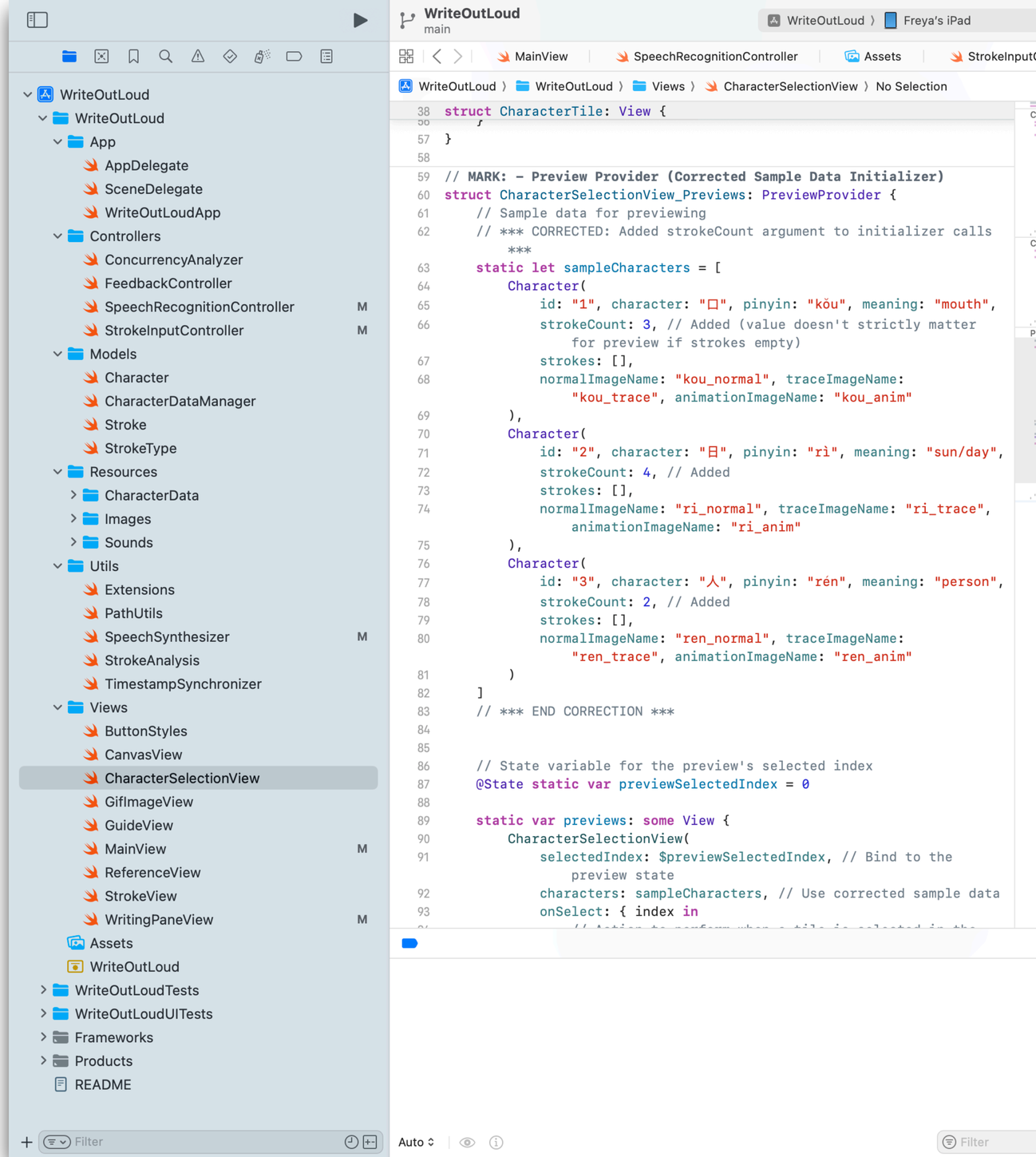
Software Development

Apple Development

- Apple Developer Program restrictions make collaboration across multiple Apple IDs difficult
- Freya has been the only code contributor 🎖️

Software Development
Code Logic

- Speech segmentation refined for accurate stroke alignment
- Homophone handling with targeted corrections
- Limited character set, manually curated for quality
- Stroke and speech sensitivity tuned through iterative testing
- **Real world data can be messy!**

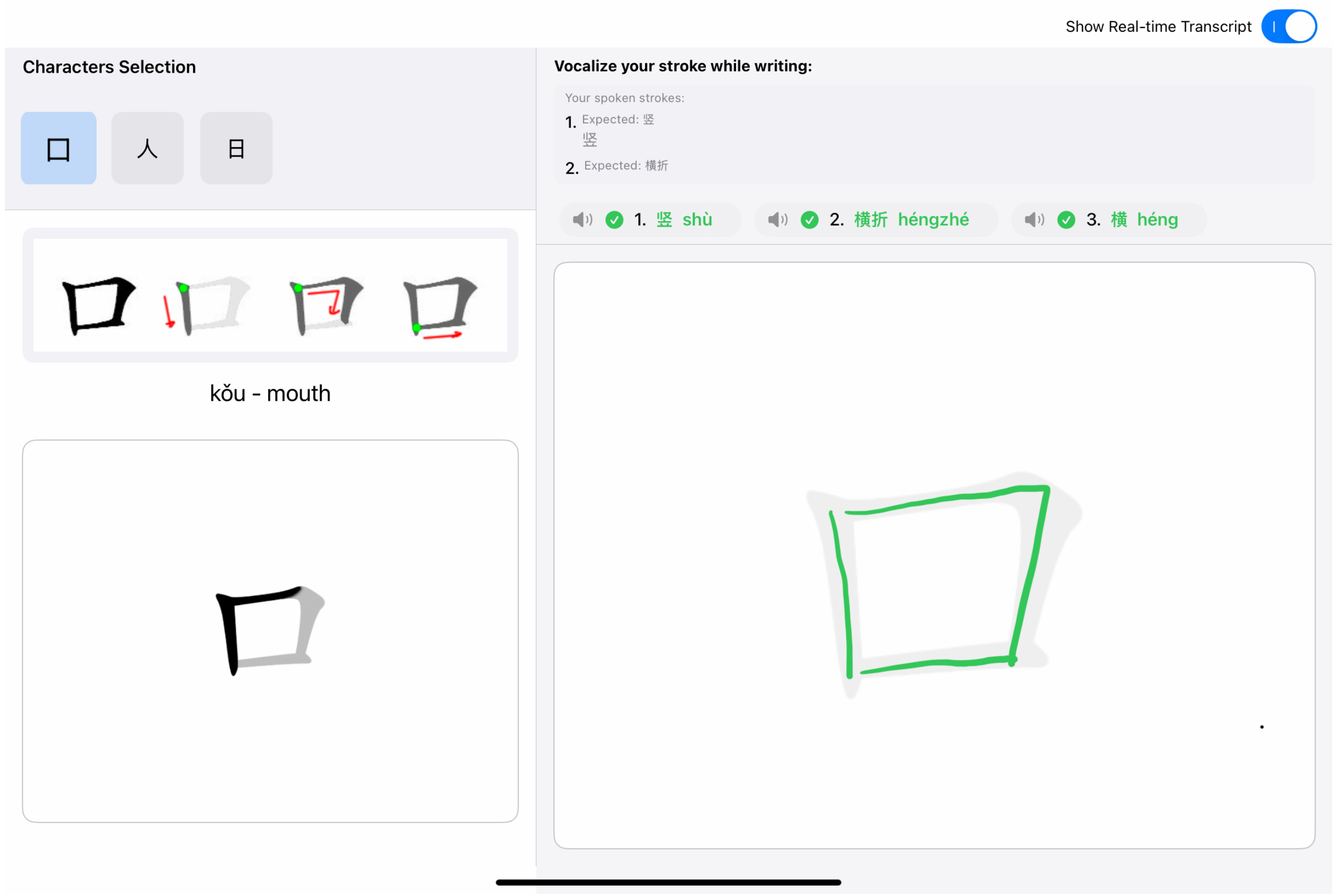


User Testing

Niche audience?

- Chinese speakers found stroke name vocalization redundant
- Non-Chinese speakers struggled with pronunciation accuracy at first
- Currently **more intuitive** for learners with some prior **pronunciation foundation**

Write Out Loud



https://github.com/junruren/Write-Out-Loud

Junru Ren <junruren@mit.edu> & Freya Tan <freya117@mit.edu>