Massive Pronunciation Training via Mobile Applications

Hui Lin
LingoChamp Inc.
h@liulishuo.com

Abstract

In this paper, we introduce a mobile application (app) that helps people practicing their English pronunciation using mobile devices. Equipped with an embedded assessment engine, the app offers accurate pronunciation assessment and feedback to a learner instantly. Moreover, game elements and mechanics are introduced to make the training experience fun, rewarding and engaging. The app turns out to be phenomenal and leads to massive pronunciation training at an unprecedented scale. Since its launch, the app has accumulated more than 20 million users. Hundreds of years of speech data are collected from more than 11 million different speakers, which is probably the largest speech corpus for Chinese spoken English in the world.

Index Terms: speech recognition, computer assisted language learning, computer assisted pronunciation training

1. The Mobile App

In this paper, we introduce our effort on mobilizing computer-assisted pronunciation training (CAPT). The goal is to not only enable accessible and ubiquitous pronunciation training on mobile devices, but also to make the pronunciation training fun such that learners are happy to keep practicing which shall then improve their ability to better speak the language eventually.

We first develop a CAPT engine that are fast, accurate and small in footprint. The footprint size matters as mobile devices are usually resource-limited. Moreover, larger app in size reduces the chance of installation on impulse through cellular networks (actually, apps that are larger than a certain size cannot be downloaded using a cellular network connection, and must instead be downloaded via WiFi). With the engine installed on device, a learner or user can practice her pronunciation with instant feedback anytime and anywhere, even without internet connections. Gamification elements are further introduced to make a learner’s experience fun, engaging, and rewarding. In particular, game mechanics like levelling up, awards, badges for achievements, feedback, challenges, are used, elevating the learning experience on our app above the mundane activities of normal pronunciation training.

Our app, called “LiuLiShuo”, was launched on the February 14th of 2013, and has since got over 20 million users. Figure 1 shows the rank history of our app in the education category of Apple’s App Store in China1. As can be seen, the app ranks top twenty most of the time.

Note that the user growth is organic and we haven’t spent a dime on promoting the app. The popularity is due to the fun and excellent user experiences it offers. Users are happy to give good reviews and recommend our app to their friends. On average, our app scores 4.5 out of 5 based on more than 12,000 ratings, and has been selected as App Store’s Best of 2013 in China.

Of course, there are many success factors in retrospect, for instance, the introduction of gamification with elements like locked lessons, challenge mode, rewards, leaderboards and etc. Above all, however, we believe the cornerstone is the accuracy and promptness of our on-device assessment engine.

The embedded assessment engine is optimized aiming to be fast, accurate, and small in footprint. To do so, several techniques are used. For example, fixed-point computation is introduced, and we use a shared Gaussian pool for all Gaussian Mixture Models in the acoustic model to speed-up the computation while reducing the size of models. As a result, the final package size of the assessment engine along with all the models is less than 10M bytes, and the engine runs smoothly even on low-end Android phones.

In terms of accuracy, the engine has been tested extensively by users of the app. Given the large number of users, it is probably the most tested pronunciation assessment algorithm ever. Users play with the engine with various ways trying to fool it. Users also listen to the recordings on leaderboards to see if rankings produced by the engine make sense. They would not keep practicing their pronunciation using the app if they thought the score given by the engine was not accurate, nor would they leave good reviews or recommend the app to others. Therefore, the popularity of the app is actually a good testimony to the accuracy of the assessment engine.

2. Acknowledgements

We thank all the colleagues at LingoChamp, including content creators, product managers, designers, iOS / Android developers, and backend engineers, without whom, such a phenomenal app would not be possible.