

JIBBIGO: SPEECH-TO-SPEECH TRANSLATION ON MOBILE DEVICES

Matthias Eck, Ian Lane, Ying Zhang, Alex Waibel

Mobile Technologies LLC, Jibbiggo

1. INTRODUCTION TO JIBBIGO

Jibbiggo is a speech-to-speech translation application for iPhone, iPod touch, and iPad devices. Jibbiggo allows the user to simply speak a sentence, and it speaks the sentence aloud in the other language, much like a personal human interpreter would. The speech-to-speech translation is bi-directional for a two way dialog between participants.

Figure 1 shows screenshots of the Chinese-English and Spanish-English versions of Jibbiggo. In addition to the spoken output, the user can read the speech recognition output and translations.



Figure 1: Jibbiggo Chinese-English and Spanish-English

The user is also able to manually correct recognition errors and just type an incorrectly recognized word.

2. AVAILABILITY

Jibbiggo has been available for the iOS platform since October 2009. Jibbiggo was initially only available for Spanish-English. 3 other language pairs were released since then with further languages under development. Jibbiggo can be used on iPhone, iPod touch and iPad with an Android version under development. Table 1 lists all language pairs and pricing.

Language Pair		Price
Spanish	English	\$24.99
Japanese	English	\$27.99
Chinese	English	\$24.99
Iraqi-Arabic	English	\$24.99

Table 1: Language availability and pricing

3. FEATURES OF JIBBIGO

The main features of Jibbiggo include:

- State-of-the-art speaker independent Speech Recognition Technology
- Vocabulary of 40,000+ words
- Statistical Machine Translation for optimal performance
- Natural voice output using SVOX speech synthesis
- Large background text dictionary

Some additional unique features of our application are described in more detail below:

3.1. No data connection required

A unique attribute of Jibbiggo is that it runs entirely on the mobile device and does not require network connectivity or remote servers to operate. Speech recognition, machine translation and text-to-speech are all performed locally on the device thus offer a significant advantage compared to network-based approaches. Most critically, users do not have to worry about potentially costly data roaming charges when using Jibbiggo in a foreign country. This is also an important advantage in disaster situations where network coverage might not be available or spotty.

3.2. Personalized vocabulary

Even though Jibbiggo is using a large recognition vocabulary of over 40,000 words, the coverage of named entities is naturally limited. It is not practical to include all the place, person, organization and product names that are relevant to all users of the system. However, named entities are critical for communication and those relevant to the current user and their location must be able to be handled by the system. To overcome this problem Jibbiggo offers the possibility for

users to add, remove and edit names to the system while it is being used in the field. Currently this is supported for first names, last names and location names.

Adding a new name to the system requires the extension of:

- Speech Recognition vocabulary in source and target language
- Machine Translation phrase table
- Text to Speech vocabulary

For speech recognition and machine translation the new name has to also be included in the language model to provide discriminative probabilities for utterances using the new word. This is supported by an internal class-based model for the respective classes.

Jibbigo can accomplish these model extensions at runtime with no delay for the user. Once the user has added a new name, he will be able to use it immediately.

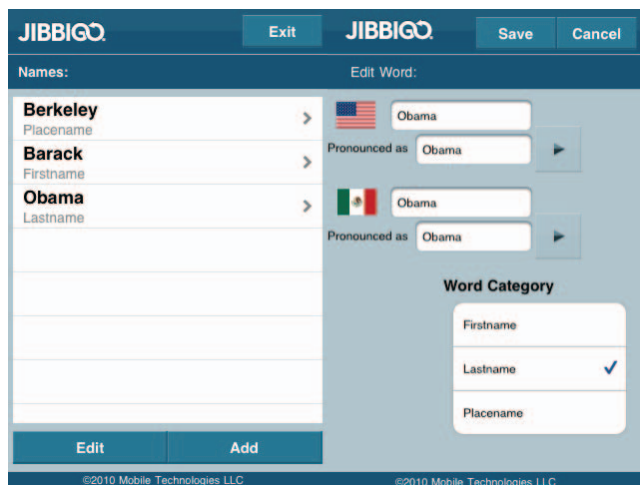


Figure 2: Adding names to Jibbigo

3.3. Error Identification and Logging

Jibbigo offers users 2 functions to log errors to Jibbigo’s servers. Individual sentences can be sent via a “thumbs down” button. The intention here is to notify Jibbigo of bad translations and low recognition performance, which will subsequently be used to improve system performance. Additionally users can also upload their whole history of sentences.

The upload consists of the spoken audio, recognition hypothesis and the translation hypothesis. Uploads are manually transcribed, translated and re-integrated to build models with higher performance. All uploaded sentences are anonymized and cannot be traced back to the individual

users. No user is actually required to upload any data so the user privacy is completely protected.

4. USAGE SCENARIOS

Tourism

The typical usage scenario is certainly a tourist visiting a foreign country. In this domain the topics of discussion will focus around hotel reservation, booking flights, sightseeing, directions, restaurants etc.

Government/Military

Another important application for a mobile speech-to-speech translation system are military operations. The military only has access to a very limited number of local interpreters at a high cost, so an automatic speech-to-speech translation system can help in many circumstances to communicate with the local population. Additionally, network connectivity is non-existent in many deployments and thus on device operation is critical.

Medical/Humanitarian

Mobile speech-to-speech translation systems are also well suited for humanitarian missions. Medical professionals often have to treat patients with very limited or no knowledge of English. This is often the case in the United States where hospitals face an increasing number of foreign speakers with whom they need to communicate with them.

In situations where translation accuracy is critical, for example in some medical and military scenarios, professional human translators are essential. The speech-to-speech translation systems, however, can support the broader translation effort by providing translation in low-tension situations, freeing up professional for more critical tasks. Additionally, in many cases where no translator is available these systems enable cross-lingual communication that would not occur.

5. MORE INFORMATION

For further information please visit www.jibbigo.com