The problems of the Rovi demo are:

a) Chanel mismatch: It uses a different recording channels to the phone (there is a channel mismatch because our models are for phone),

b) Frequency sampling mismatch: Rovi uses 16khz and our system uses 8khz (sampling also for analog phone),

c) Accent mismatch: The user do not use native pronunciation USA.

All this mismatches make the system works poorly.

We have made a planning aiming reduce this mismatches. This planning will be evolving along the project to adapt it to the results we will obtain (we can cancel a non-promising line of working or too costly in time).

In very short term there are few things we can do, to improve the Rovi demo, which we have not done yet. Options:

1) Find some rules of pronunciation of English USA by an Indian [MT]. With these rules, retrain the phonetic transcriber [JP]. Acoustic models won't be touched. We would adapt to the person to be evaluated. (2015/01/19)

2) Retrain the acoustic models with captions of some Indian international television channel [MC]. (2015/01/19)

3) Remove the garbage acoustic model from the BNF of the demo and write in this grammar all standard phrases that we think the person can say (improve FILLER words) [AM] [MT]. (2015/02/02)

In the short term. Options:

1) Search transcription rules to create alternative dictionaries for common pronunciations in USA [MT]. (2015/02/02)

2) Get free databases with non-native voices USA to train the acoustic models [MC] [AM] [HC].(2015/02/02)

3) Insert acoustic models from other languages. [MC] [JP]. (2015/02/02)

4) Improve our system for speaker adaptation / retrainning for working with less signal. We will have to talk a while before trying the demo. The challenge is to keep it under 5 minutes [JA] [AN] [JP]. (2015/02/02)

In the medium term. Options:

1) Doing robust channel adaptation of the acoustic models [CG] [JA] [JP]. (2015/03/09)

2) Train models Microphone (16khz) for native and non-native USA. [MC] [JP] [AN]. (2015/04/06)