



An iPhone App to Help Travelers Find their Way in the Tokyo Region

Melvyn Hunt & Yoon Kim



Contents



- Overview of the application
- Server-based or embedded ASR?
- User Behavior
- Accuracy and Geographical Coverage
 - bootstrap training
 - increased geographical coverage
- Public reaction
- What we gained from this development

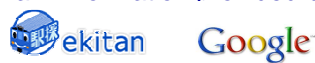
Overview of the Application

- Most travel in the *Kanto* region of Japan is by train/subway
 - 1,919 stations in the Kanto region
- *iPhone* now very popular in Japan
- Novauris invited to develop an iPhone app
 - user is asked to say:
 - “A ka-ra B ma-dey” (*from A to B*)
 - e.g. “Shinjuku ka-ra Yokohama ma-dey”
 - and gets back information on the journey from station A to station B
- App is free to the user, supported by advertising.



“Koetan!”

- **Koetan (コエタン)**
= “voice search”
- Voice search of train routes
 - Launched in September 2009
 - Response provides:
 - list of trains and changes
 - the fare
 - route map
 - map of area around the destination
 - route by car or foot to the departure station (Google Maps)
 - Train information provided by



Server-based or Embedded ASR?

- ❑ Koetan is server-based
- ❑ We implemented a fully embedded version on the iPhone and it worked well
- ❑ The embedded version has two big advantages:
 - it's faster (no network delays)
 - it will still work where there's no network signal
- ❑ So why not use the embedded version?



Why not use embedded ASR?

- ❑ Can't easily use advertising for the revenue
 - so would have to charge for the app
- ❑ Initial download would take longer due to app size
 - so that makes two barriers to initial uptake
- ❑ Network coverage is excellent in the Tokyo area
- ❑ But actually the **BIG REASONS** are that server-based operation lets us :
 - observe user behavior
 - gather large amounts of speech to use for training
 - update the ASR and the database quickly and often

Observations of User Behavior

- Japanese users are like Western users:
 - they don't follow instructions carefully
 - they use a minimum number of words
- Asked to say: "A ka-ra B ma-dey" (*from A to B*)
 - about 40% leave out "ma-dey" (*to*)
 - about 10% leave out "ka-ra" (*from*)
 - about 10% say only "B" (the destination station)!
 - about 1% add "eki" (*station*) to the station names
- Users sometimes ask for major stations outside the specified region
- Around 15% of inputs appear to be to play with the ASR

Adapting the App to User Behavior

- ❑ A probabilistic grammar allowed "ma-dey" and "ka-ra" to be omitted and "eki" to be inserted with penalties reflecting frequency of occurrence.
 - making these items optional without probabilities actually harmed accuracy
 - but with them it noticeably helped
- ❑ The grammar was extended to be compatible with inputs of only the destination station
- ❑ Major stations outside the region (*e.g.* Kyoto, Osaka) were added

Accuracy and Geographical Coverage

- ❑ Service began with coverage of the Tokyo Prefecture (central Tokyo) only
 - and with undertrained speech modelling
- ❑ Users perceived the accuracy to be good
 - despite calls typically being from noisy stations or noisy underground trains
 - but many requested stations were outside the area
- ❑ Area extended first to Greater Tokyo (4 prefectures), then to Kanto (7 prefectures, 1/3 of Japan population)
- ❑ Training on speech from real users more than compensated for the harder task (160k → 3.7M routes)
- ❑ Accuracy >95% (correct journey displayed)
 - for valid inputs



Summary Statistics

	% of all inputs	% of valid inputs
Defective inputs		
Empty input	5	-
Unclear or out-of-grammar stations	15	-
Extra talking in speech	1	-
Too many stations / stations repeated	1	-
Only one station spoken	9	-
Total	31	-
Form of input		
“eki” appended to station names	1	1
A B only (no “ka-ra” or “ma-dey”)	9	13
A ka-ra B (no “ma-dey”)	41	59
Proportion Correctly Recognized	66	96
User Response		
User goes on to get train information	50	73
Response correct but user stops	15	22





Public Reaction



- ❑ Apple Japan made Koetan a “Featured App”
- ❑ Listed in top 5 location-related iPhone apps for Japan
- ❑ Much positive coverage:
 - AppBank
 - CNET Japan
 - 6 other positive reviews in Japanese; 4 in English
- ❑ User web feedback comments have been ~100% positive on the ASR
 - but critical of the usefulness of the app



What we Gained from Koetan

- ❑ Only modest revenue so far
- ❑ So was it worth it?
 - You betcha! 
 - Because:
 - We gained a large speech corpus from real users
 - we know of no publicly available Japanese speech corpus for ASR for which commercial use is allowed
 - We gained credibility in Japan
 - leading to major, as yet unannounced, agreements
 - We got huge personal satisfaction
 - from developing ASR in a language we don’t speak and seeing Japanese speakers being impressed!



novauris

Speak, *Find.* ...EASY!

More demo Videos on  **YouTube**
Broadcast Yourself™

www.youtube.com/novauristechnologies

