

Japanese Dictation Toolkit

– Free Software Repository for Speech Recognition –

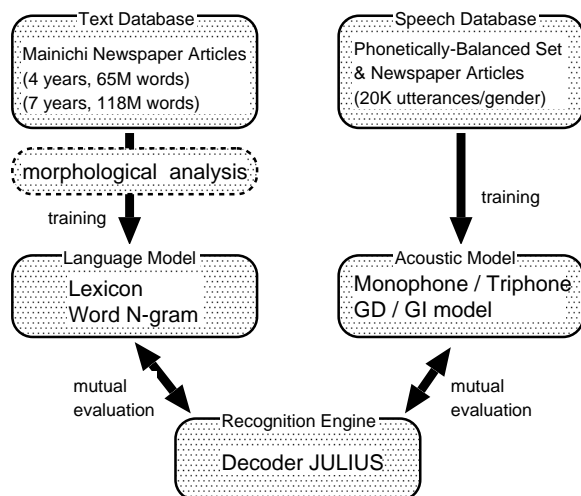
Tatsuya Kawahara (Kyoto Univ.), Tetsunori Kobayashi (Waseda Univ.),
Kazuya Takeda (Nagoya Univ.), Nobuaki Minematsu (Toyohashi Univ. of Tech.),
Katsunobu Itou (ETL), Mikio Yamamoto (Tsukuba Univ.),

Atsushi Yamada (ASTEM), Takehito Utsuro, Kiyohiro Shikano (Nara Inst. of Sci. & Tech.)

<http://winnie.kuis.kyoto-u.ac.jp/pub/julius/doc/00readme.html>

1. Overview

- A sharable software repository
- Project sponsored by a governmental organization IPA (Information-technology Promotion Agency), Japan
- Collaboration of researchers of different academic institutes in Japan
- Available to public freely ¹



2. Specification of Models and Programs

2.1 Acoustic Model

Continuous density HMM (HTK format) trained with ASJ (Acoustical Society of Japan) databases

model	#states	#mixtures	gender
monophone	129	4, 8, 16	GD, GI
triphone 1000	1000	4, 8, 16	GD
triphone 2000	2000	4, 8, 16	GD, GI
triphone 3000	3000	4, 8, 16	GD

¹For further information:

<http://www.lang.astem.or.jp/dictation-tk/>
mailto: dictation-tk-request@astem.or.jp

2.2 Japanese Text Processor

- Morphological Analyzer: ChaSen
- Pronunciation Post-Processor: ChaWan to handle undelimited Kanji-Kana texts

2.3 Lexicon and Language Model

vocabulary size	coverage
5000	88.2%
20000	96.5%

Word 2-gram, 3-gram (ARPA format) trained with Mainichi newspaper articles of 7 years

2.4 Decoder: Julius

	acoustic model	language model	search approx.
1st pass	intra-word CD	2-gram	1-best
2nd pass	inter-word CD	3-gram	N-best

3. Key-Property of Toolkit

- open format & interface
- easy to revise and replace modules
- reasonable performance as the baseline word accuracy 85~93% / 1~8 real time
- current 1998 version
20K model & system
Unix platform
- condition
Copyright reserved
NO WARRANTY

REFERENCES

- [1] T.Kawahara et al. Sharable software repository for Japanese large vocabulary continuous speech recognition. In *Proc. ICSLP*, pp.3257–3260, 1998.
- [2] K.Itou et al. The design of the newspaper-based Japanese large vocabulary continuous speech recognition corpus. In *Proc. ICSLP*, pp.3261-3264, 1998.

