1) Introduction

- Frisian Language
  - regional official language of the Netherlands
  - mostly spoken in the province of Fryslân
  - approximately half a million speakers
  - most speakers are bilingual due to the extensive influence of Dutch
  - code-switching is common practice in daily conversations
  - Member of West Germanic language family
    - closely related with English and Dutch

- Linguistically well-researched, however few speech and language technology applications are available
  - Frisian speech synthesizer, Google Translate...
  - Omrop Fryslân
    - regional public broadcaster with a radio station
    - a TV channel both broadcasting in Frisian

2) FAME! Project

- Disclose the Omrop Fryslân archives containing recordings from 1950s
  - Develop a user-friendly search interface for spoken documents from Omrop Fryslân archives with more than 2600 hours of radio broadcasts
- Relevant applications towards building this spoken document retrieval system:
  1. Automatic speech recognition
  2. Speaker identification
  3. Flexible search interface
  4. Project Partners:

3) Basic Frisian ASR system

- Challenges
  - Low resources available
  - Code-switching nature of Frisian
  - Complex vowel system

- Acoustic Model
  - FAME! Database is created
  - Frysk Akademie text corpus
  - Omrop Fryslân news articles
  - Frisian Wikipedia
  - Transcriptions of training speech
  - Fluerjy Lexicon
  - Elx Dutch Lexicon

- Language Model

- Lexicon

4) FAME! Frisian Radio Broadcast Database

- Preparation
  - Manually annotating the radio broadcasts from Omrop Fryslân
  - Collaboration with Frysk Akademie
  - Annotations include orthographic transcription, speaker ids, spoken language, code-switching details, dialect info
  - A modified annotation protocol has been created

- Some statistics
  - 18.5 hours of radio broadcasts annotated in total
  - Longitudinal data: recordings from 1966 to 2015
  - More than 500 speakers, 309 with known identity
  - 21 speakers appear at least 3 times

- 3939 code-switching cases:
  - 2896 cases: Frisian speaker switches to Dutch
  - 95 cases: Dutch speaker switches to Frisian
  - 848 cases: Speakers use a mixed-word that is neither Frisian nor Dutch

5) Frisian Language Model and Lexicon

- Language Model
  - Frisian text corpus: ~2,375,000 sentences
  - Training speech transcription: ~1,750,000 sentences
  - Dutch text corpus (CGN): ~580,000 sentences
  - Monolingual and bilingual N-gram models are trained

- Lexicon
  - Complete Frisian lexicon: ~340k words
  - Complete Dutch lexicon: ~1.1M words
  - Frisian phonetic alphabet contains 2 consonants, 20 monophones, 16 falling diphongs, 8 rising diphongs and 6 triphongs
  - For bilingual lexicon, Dutch phones are mapped to the phonetically closest Frisian phone.
  - Grapheme-to-Phoneme (G2P) models are learned to handle the out-of-vocabulary (OOV) words in training data

6) Initial Recognition Experiments

- Data from Frisian speakers
  - FAME! Database is divided into three parts
    - Training set: 8h 20m
    - Development set: 1h
    - Test: 1h
  - Acoustic models (AM)
    - KALDI speech recognition toolkit is used
    - GMM-HMM and subspace GMM (SGMM) are trained on LDA-MLLT features
    - Speaker adapted training (SAT); FMLLR-adapted features
  - Language models (LM)
    - 3-gram interpolated modified Kneser-Ney
    - Omrop Fryslân and bilingual LM are compared
  - Lexicon
    - Frisian lexicon contains ~95k words
    - Bilingual lexicon contains ~150k words
    - Various phonetic alphabets are compared:
      - mono: cons. + monoph. + fall. monoph + fall diph.
      - rise: monoph + rise diph.
      - diph: monoph + all diph.
      - triph: monoph + all triph.
- Word error rates (WER) in % on the development set:

7) Results (I) – Phonetic Alphabet

- The choice of the phonetic alphabet has a minor effect on the recognition accuracy
  - Superior performance of dtrn is explained by the limited amount of training data
  - In the following experiments, mono is adopted

8) Results (II) – AM, LM and Lexicon

- WERs using mono- and bilingual lexicon and LM
  - WERs with G2P for OOV words (Best results for SI system)

9) Conclusion

- Initial recognition results are promising for an accurate spoken document retrieval system
- Future work: Investigating deep architectures and recognition schemes with flexible lexicon for code-switching ASR

---

[Image: Radboud University Radboud Universiteit]