1) Introduction

3) Basic Frisian ASR system

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

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

4) FAME! Frisian Radio Broadcast Database

- Preparation
  - Manually annotating the radio broadcasts from Omrop Fryslân
  - Collaboration with Frysk Akademy
  - Annotations include orthographic transcription, speaker ids, spoken language, code-switching details, dialect info
- 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: ~13,750 sentences
  - Dutch text corpus (CGM): ~580,000 sentences
  - Monolingual and bilingual N-gram models are trained
- Lexicon
  - Complete Frisian lexicon: ~340k words
  - Most speakers are Fryske (Fryslân)
  - Transcriptions to the Fryske Omrop public broadcaster with a radio station
  - Regional public broadcaster with a radio station and a TV channel both broadcasting in Frisian
  - Speaker adapted training (SAT): FMLLR
  - Linguistically well-researched, however few speech and language technology applications are available
  - Frisian speech synthesizer, Google Translate...

6) Initial Recognition Experiments

- Speech data from Frisian speakers
  - FAME! Database is divided into three parts
  - Training set: 8h 20m
  - Development set: 1h
  - Test set: 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
  - Frisian LM 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: mono + fall diph.
    - Rise: mono + rise diph.
    - Dipth: mono + all diph.
    - Triph: mono + all triph.
  - Word error rates (WER) in % on the development set

7) Results (I) – Phonetic Alphabet

<table>
<thead>
<tr>
<th></th>
<th>GMM</th>
<th>SGMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>mono</td>
<td>50.65</td>
<td>44.56</td>
</tr>
<tr>
<td>fall</td>
<td>50.34</td>
<td>45.01</td>
</tr>
<tr>
<td>rise</td>
<td>50.85</td>
<td>45.31</td>
</tr>
<tr>
<td>diph</td>
<td>50.99</td>
<td>45.10</td>
</tr>
<tr>
<td>triph</td>
<td>50.23</td>
<td>44.88</td>
</tr>
<tr>
<td>dtriph</td>
<td>50.82</td>
<td>45.46</td>
</tr>
</tbody>
</table>

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

8) Results (II) – AM, LM and Lexicon

<table>
<thead>
<tr>
<th></th>
<th>Devel</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lex</td>
<td>GMM</td>
<td>SGMM</td>
</tr>
<tr>
<td>FR</td>
<td>FR</td>
<td>FR</td>
</tr>
<tr>
<td>NL</td>
<td>NL</td>
<td>NL</td>
</tr>
<tr>
<td>Fryslân</td>
<td>Fryslân</td>
<td>Fryslân</td>
</tr>
</tbody>
</table>

- 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