The following full text is a preprint version which may differ from the publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/157399

Please be advised that this information was generated on 2019-05-13 and may be subject to change.
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

**Phonology**
- Frisian consonants are more or less similar to Dutch consonants
- Frisian has more vowels (falling and rising diphthongs, triphthongs, rassilation)

**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 and a TV channel both broadcasting in Frisian

3) Basic Frisian ASR System

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

**FAME! Database is created**
- Frysk Akademy text corpus
- Omrop Fryslân news articles
- Frisian Wikipedia
- Transcriptions of training speech
- Fluenty Frisian Lexicon
- Elex Dutch Lexicon

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
- A modified annotation protocol has been created

**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 switch 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
- Complete Dutch lexicon: ~1.1M words

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 (SGM) are trained on LDA-MLLT features
- Speaker adapted training (SAT): FMLLR-adapted features

**Language models (LM)**
- 2-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 + fll dipth.
  - rise: mono + rise dipth. + diph: mono + all dipth.
  - trig: mono + all trig. + diph: all dipth.

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>55.34</td>
<td>45.01</td>
</tr>
<tr>
<td>rise</td>
<td>56.85</td>
<td>45.31</td>
</tr>
<tr>
<td>diph</td>
<td>56.09</td>
<td>45.10</td>
</tr>
<tr>
<td>riph</td>
<td>50.23</td>
<td>44.88</td>
</tr>
<tr>
<td>dtrin</td>
<td>50.82</td>
<td>45.46</td>
</tr>
</tbody>
</table>

- Word error rates (WER) in % on the development set -

- The choice of the phonetic alphabet has a minor effect on the recognition accuracy
- Inferior performance of dtrin 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><strong>Lex</strong></td>
<td><strong>LM</strong></td>
<td><strong>GMM</strong></td>
</tr>
<tr>
<td>FR-NL</td>
<td>49.20</td>
<td>45.85</td>
</tr>
<tr>
<td>FR-NL</td>
<td>49.39</td>
<td>45.44</td>
</tr>
<tr>
<td>FR-NL</td>
<td>49.16</td>
<td>43.49</td>
</tr>
</tbody>
</table>

- **WERs using mono- and bilingual lexicon and LM -**

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