**EXPERIMENTS AND RESULTS**

**Experimental protocol**
- **DECODA project corpus** of conversations:
  - Train = 740 / Dev = 175 / Test = 392
  - Automatic Speech Recognition (ASR) system: Speeral
  - Word Error Rate (WER) with stop-list of 126 words:
    - train = 33.8% / dev = 45.2% / test = 49.5%
- **8 conversation themes**
- **TRS**: manual transcriptions - ASR: automatic transcriptions
- **3,000 hidden topic spaces** with a different topic number was built using the train corpus

**Theme hypothesization accuracies using different c-vectors and GMM-UBM sizes**

<table>
<thead>
<tr>
<th>c-vector size</th>
<th>DEV</th>
<th>TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>32</td>
<td>64</td>
</tr>
<tr>
<td>80</td>
<td>88.8</td>
<td>86.5</td>
</tr>
<tr>
<td>100</td>
<td>91.2</td>
<td>92.4</td>
</tr>
<tr>
<td>120</td>
<td>89.5</td>
<td>92.2</td>
</tr>
</tbody>
</table>

- **Classification performance is stable** (5.9 points difference for dev)
- Using comparable training and testing configurations allows to achieve the best classification performance

**CONCLUSION**
- In spite of very high WER, possible to classify effectively documents with the proposed compact representation (c-vector) with an accuracy of 85% + allows us to both solve the difficult choice of the right number of topics and the theme proximity
- **Future work** will seek to find the best combination of LDA hyper-parameters and evaluate effectiveness in other NLP tasks