LEXICAL ACCESS-BASED CONFIDENCE MEASURE FOR A SPANISH KEYWORD SPOTTING SYSTEM
Javier Tejedor, Dong Wang, Simon King and Jose Colas
e-mail: Javier.tejedor@uam.es
CSTR-University of Edinburgh

ABSTRACT
Keyword spotting deals with the search of a reduced set of keywords in audio content. Phone Lattice-based approaches are very fast but achieve poor results. HMM-based keyword spotting systems deal with filler models to absorb the Out-of-vocabulary (OOV) words and achieve best results although they are slower. We propose a technique which combines them in order to perform a confidence measure to reduce the false alarm rate achieved in the HMM-based keyword spotting module over a Spanish Keyword Spotting system. Different filler models are investigated within the experiments.

SYSTEM ARCHITECTURE

Keywords proposed in the Keyword spotting module are accepted or rejected by the confidence measure module based on a lexical access which retrieves the keyword which best matches with the sequence of phones retrieved by the Phonetic decoding. The confidence measure is based on this keyword and on a threshold estimated from the sequence of phones of the correct keywords retrieved by the Keyword Spotting module.

EXPERIMENTS
80 keywords in the test geographic corpus in Albayzin database.
Filler models: Phones (AM), phonemes (PM), broad class (BM) and unique model (UM).
39 MFCC's (13 MFCC's + first and second derivatives).
GMM/HMM Context-independent models from the training phonetic corpus.
Pseudo 6-gram (Kim et. al., 2004) as language model in HMM-based keyword spotting process.

<table>
<thead>
<tr>
<th>Modelling</th>
<th>#KW</th>
<th>#Occurrences</th>
<th>#Hits</th>
<th>#False Alarms</th>
<th>FOM</th>
</tr>
</thead>
<tbody>
<tr>
<td>AM</td>
<td>80</td>
<td>1672</td>
<td>1073</td>
<td>140</td>
<td>78.04</td>
</tr>
<tr>
<td>PM</td>
<td>80</td>
<td>1672</td>
<td>1162</td>
<td>171</td>
<td>78.85</td>
</tr>
<tr>
<td>BM</td>
<td>80</td>
<td>1672</td>
<td>1216</td>
<td>211</td>
<td>76.17</td>
</tr>
<tr>
<td>UM</td>
<td>80</td>
<td>1672</td>
<td>1195</td>
<td>211</td>
<td>73.87</td>
</tr>
</tbody>
</table>

CONCLUSIONS
Lexical decoder allows us to reduce the high number of false alarms achieved in the Keyword Spotting module: Hybrid architecture.
Future work: Search in lattices.

REFERENCES