

# The European SQALE Project (1993-1995)

- ARPA large vocabulary evaluation paradigm adapted for European languages (British English, French and German)
- Separate recognition systems for each language. Different sites built systems for 2 or 3 of the languages.
- Overall conclusion: The same general approach to recognition system design is applicable to all languages, some language specific problems. (homophones in French, compounding in German)
- Also attempted cross language comparisons but found this was impossible.

# Multilingual speech recognition in seven languages - Ulla Uebler (2001)

- Three different approaches:
  - portation - same system, language specific training data
  - cross-lingual - data from more than one language for training, testing on unseen language
  - simultaneous - one recognizer for all languages involved
- Different schemes for mapping phones from one language to another (native, phonetic, data-driven)
- Conclusions:
  - some languages are easier for ASR
  - languages need to be closely related for sharing to be effective
  - sharing acoustic information beneficial, but not language models

# Language-independent and language-adaptive acoustic modeling for speech recognition (Schultz & Waibel, 2001)

- Varied source languages - limited target data
- GlobalPhone database (10 languages), global unit set
- Monolingual systems: WERs vary greatly (scripts, letter-to-sound relation, sound system, segmentation, morphology)
- Multilingual systems: always degradation compared to monolingual
- Adaptation:
  - multilingual initial models outperform monolingual initial models
  - complementary set of languages is more important than number of languages.
  - polyphone decision tree specialization procedure - d-tree is regrown according to limited adaptation data in target language

# Multilingual acoustic models - sharing acoustic information

## - beneficial or not?

- Number of different source languages, some data for target language.
- Recognition of target data: better, same or worse than monolingual recognizer?
  - better: Niesler (2006), Nieuwoudt & Botha (2002)
  - same: Köhler (1998)(using MAP adaptation)
  - worse: Schultz & Waibel (1998), Uebler (2001)
- Again, it seems languages need to be closely related to benefit.

# Units of sharing information between languages

- global unit set = phonemes (Schultz et al., Köhler, Uebler, Nieuwoudt & Botha etc.)
- articulatory features in addition to phonemes (Stüker et al.)
- articulatory features (Chang et al.)
- graphemes (Killer et al.)