Supporting cultural diversity: Expanding the language coverage of the ETSI spoken command vocabulary standard.

Mike Tate, Helge Hüttenrauch, Martin Boecker, Françoise Petersen, Rosemary Orr,
Please put in the appropriate header with names, if so desired.

Schneider-Hufschmidt Matthias; 06/09/2005
Need for Standardisation

- No previous standard - industry or other bodies
- Increased prevalence of speech systems - therefore transfer of learning important
- With enlarged EU comes increased national identity - therefore need to expand current standard
- Facilitates e-inclusion and a multicultural Europe
- Local language ASR systems increases local use - prevalence of English is a barrier
- Greater uptake and use of ICTs - good for consumer QoL, service providers and equipment makers
Advantages of Speech

- Fundamental Human Paradigm, natural
- Useful for people with limited access to terminals and displays
- Facilitates e-inclusion and narrows digital divide
- Safe
- ASR technology commercially available
- Becoming cheaper and more reliable
ETSI STF 326 approach

- Definition of languages covered
- Definition of commands covered
- Empirical data collection
- Consultation and publication
Generic spoken command vocabulary for ICT devices and services

ETSI STF 182

- The ETSI Standard (ES) 202 076 specifies user tested spoken commands for five languages
  - French
  - German
  - Italian
  - Spanish
  - English
Generic spoken command vocabulary for ICT devices and services

ETSI STF 326

Expansion of Standard to cover 30 Languages
- Expansion of EU
- Strong presence of Russia
- EFTA Countries - Switzerland, Norway, Liechtenstein, Iceland
Current Voice Commands

- **Definition of commands covered:**
  - **Context-independent common commands**
    (e.g. Options, Main menu, Standby)
  - **Context-dependent common commands**
    (e.g. Help, Repeat)
  - **Core commands**
    (e.g. Yes, No)
  - **Digits**
  - **Name and digit dialling**
    (e.g. Dial, Call)
  - **Basic call handling and supplementary services**
    (e.g. Answer, Busy)
Current Voice Commands

- Definition of commands covered:
  - Media control
    (e.g. Play, Pause)
  - Browseable list for navigation
    (e.g. Next, Previous)
  - Editing commands
    (e.g. Edit, Delete)
  - Device setting
    (e.g. Choose network, Sound off)
  - Word spotting mode
    (e.g. Wake up)
## Example Table

### Table 5: Name and digit dialling

<table>
<thead>
<tr>
<th>Index</th>
<th>ICT device/service function</th>
<th>English spoken command</th>
<th>French spoken command</th>
<th>German spoken command</th>
<th>Italian spoken command</th>
<th>Spanish spoken command</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 (see note 1)</td>
<td>Initiate digit dialling sequence</td>
<td>Dial</td>
<td>Composer</td>
<td>Wählen</td>
<td>Componi</td>
<td>Marcar</td>
<td>Initiate a call to a number</td>
</tr>
<tr>
<td>5.2 (see note 1)</td>
<td>Dial a number or name</td>
<td>Call</td>
<td>Appeler</td>
<td>Verbinden mit</td>
<td>Chiama</td>
<td>Llamar</td>
<td>Initiate a call to a name or number</td>
</tr>
<tr>
<td>5.3</td>
<td>Home phone number (location)</td>
<td>Home</td>
<td>Maison</td>
<td>Privat</td>
<td>Casa</td>
<td>Casa</td>
<td>Call the stored home number</td>
</tr>
<tr>
<td>5.4</td>
<td>Work phone number (location)</td>
<td>Work</td>
<td>Travail, Bureau</td>
<td>Büro, Arbeit</td>
<td>Ufficio, Lavoro</td>
<td>Trabajo</td>
<td>Call the stored work number</td>
</tr>
<tr>
<td>5.5</td>
<td>Mobile phone number (location)</td>
<td>Mobile</td>
<td>Mobile, Portable</td>
<td>Mobil, Handy</td>
<td>Cellulare</td>
<td>Móvil</td>
<td>Call the stored mobile number</td>
</tr>
<tr>
<td>5.6</td>
<td>Car phone number (location)</td>
<td>Car</td>
<td>Voiture</td>
<td>Auto</td>
<td>Auto</td>
<td>Coche</td>
<td>Call the stored car number</td>
</tr>
<tr>
<td>5.7</td>
<td>Personal</td>
<td>Personal</td>
<td>Numéro</td>
<td>Eigene</td>
<td>Numero</td>
<td>Número</td>
<td>As above, but</td>
</tr>
</tbody>
</table>
Three phases of Data Collection and Analysis

- Phase 1 - Elicitation of command words
- Phase 2 - Validation of commands words
- Phase 3 - Phonetic discrimination and subsequent ranking of commands and alternative commands
Methodology

1. Spontaneous generation of potential command words
   - Textual descriptions.
   - Questionnaire vs Interview method

2. Confidence rating of command words
   - User testing of alternative commands

3. Phonetic discrimination
   - Recognizer field test
   - Pronunciation dictionary test
Data Collection

- Previous method - Online data collection, however....
- Increase in languages from 5 to 30, and....
- Many states are not active ETSI members - no formal links, and....
- Many have low penetration of PC ownership and/or internet access, therefore....
- Interview method
Data Collection - how do we do this with 30 Languages?

- Identify International Centre - University College Utrecht
- Recruit Interviewers - Native Language Assistants
- Ability to contact interviewees in their native countries
- Familiarity with research methods
- Special Induction Session - standardisation, functions and CWDs, interviewing, overcoming problems,…
Phase 1 - Spontaneous generation of potential command words

- Textual descriptions method
  - Carefully worded description of functions
  - Descriptions sometimes verbose.
  - Participant responds with appropriate one-word description.
Carefully Worded Descriptions (CWD)

- You have told ‘Speak-to-me’ who you want to be connected with, but there are three phone numbers to choose from for that person. What command would you use to be connected to where the person lives?

- You have just used ‘Speak-to-me’ to make a phone call to a number you didn’t write down. Now you want to be connected to that number again.

- You don't know what ‘Speak-to-me’ expects you to do next, so you want guidance from it.
Phase 2 - Confidence rating of command words

- Histogram frequency of words from initial data capture
- Confidence rating from different participants on each alternative command name
- Measured on a 5-point scale or...
- Choose the command word for which they have most confidence.
Phase 3- Accoustic Discrimination

- Optimal discrimination between commands

  - Recogniser field test with ASR systems...but sensitive to
    - Different types of recognition system
    - Test database (recorded, live, styles)
    - 30 different recognisers required

  THEREFORE....

- Pronunciation dictionary test
  - Accoustic analysis of command words
  - Using lookup table of phone sequences
Phonetic Discrimination

- **Preparation**
  - Recordings of words
  - Semi-anechoic chamber
  - IPA transcriptions

- **Confusability testing**
  - Using number of phones in competing words
  - Phonetic confusability
  - Only for words simultaneously available to the user
Phonetic discriminability

- ‘STOP’ - 4 phones
- START - 5 phones

But

- ‘STOP’ - 2 Different phones
- START - 3 Different phones

In contrast

- Begin - 5 phones
- End - 3 phones

And no common phones
Selecting the Command Set

- Confidence Rating
- Confusability Testing
- Judgement (Common Sense)
Timetable

- Start of the work: March 2007
- Initial Draft ETSI Standard: September 2007
- Interim report to EC/EFTA: December 2007
- Final Draft ETSI Standard: September 2008
- Publication of ETSI Standard: January 2009
Thank You!

Mike Tate
Chimera
University of Essex

mtate@essex.ac.uk
Expansion of Standard to cover 28 Languages

- Bulgarian, Czech, Croatian, Danish, Dutch, English, Estonian, Finnish, French, Gaelic, German, Greek, Hungarian, Icelandic, Italian, Latvian, Lithuanian, Macedonian, Maltese, Norwegian, Polish, Portuguese, Romanian, Romansch, Russian, Slovak, Slovenian, Spanish, Swedish and Turkish.
LIST COMMANDS
Diffusion processes

Stereotype of when ‘the elderly’ adopt

- Innovators, technology enthusiasts
- Early adopters, visionaries
- Early majority pragmatists
- Late majority conservatives
- Laggards, skeptics

Customers want technology and performance
Customers want solutions and convenience

Relative % of customers

Time