Engineering a Quality User Experience for Mobile Speech Enterprise Applications

Anna M. Wichansky, Ph.D CPE
Senior Director, Applications UX
The Business Problem:

Sales Reps don’t like using CRM systems.

So often they don’t, and Managers don’t get full benefits of CRM systems.
Introducing Oracle Voice

Speech to Text and Text to Speech conversion
The Usability Engineering Problem
Business Requirements

• Acquire commercial license for speech recognition and text-to-speech
  – Enterprise domain vocabulary, proper names, industry jargon
  – Multiple mobile platforms usage
  – Multiple deployments (Oracle sales force, various customers)
  – Multiple languages
  – Multiple application platforms
  – Wide variety of usage environments and contexts
  – Potentially millions of users, hundreds of job profiles
  – SLA: flawless operations, high accuracy, fast performance
Factors to consider:

- Features and functions
- Expected usage patterns
- Speech recognition accuracy
- Performance optimization
- Languages, dialects, accents
- Security of customer data
- Best practices
What application features and functions to support?

Context of Use
Task Analysis
User Profile
Development

Usage Patterns
User Requirements
Technology Requirements
Usage Patterns

• How much will they use the product?
  – Vendor royalties
  – Support
  – SLA
  – Pricing
  – Customization
  – Languages
  – Performance

• Usability Lab Studies

• Sales Field Studies
Speech recognition accuracy

- Standard corpus development for testing
- SDK out of the box vs. with Oracle app
- Measuring through phone microphone
- Speech expert evaluations
- Vendor comparisons
- Adding vocabulary to existing toolkits
- Natural language processing
- NIST SCLITE method: WER
- Final acceptance testing
Acceptance Testing Set-up

- Laptop playing recordings separated by 5 seconds of pause
- Monitor
- Attenuator to reduce sound level
- Special cable for handset MIC. input
- Speech product Exerciser
Performance criteria

- How will speech technology work under production conditions?
- Delay in text response after speech input?
- Delay in speech output after text input?
- Speech on server vs. on smartphone?
- Simulating user loading conditions?
- App crashes?
- Data loss?
### Languages, dialects, accents

<table>
<thead>
<tr>
<th>ORCL Language</th>
<th>ORCL Code</th>
<th>Vendor Code/Lang1</th>
<th>Vendor Code/Lang2</th>
<th>Vendor Code/Lang3</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabic</td>
<td>ar</td>
<td>ara-EGY</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chinese (Simp)</td>
<td>zh-CN</td>
<td>cmn-CHN (Mandarin)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chinese (Trad)</td>
<td>zh-TW</td>
<td>cmn-TWN(Taiwan-Mandarin)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dutch</td>
<td>ne</td>
<td>nld-NLD/Dutch</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>French</td>
<td>fr</td>
<td>fra-FRA (French(EU))</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>German</td>
<td>de</td>
<td>deu-DEU (German)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Italian</td>
<td>it</td>
<td>ita-ITA (Italian)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Japanese</td>
<td>ja</td>
<td>jpn-JPN (Japanese)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Korean</td>
<td>ko</td>
<td>Kor-KOR (Korean)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Portug-BR</td>
<td>pt-BR</td>
<td>por-BRA (Portuguese-BR)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spanish</td>
<td>es</td>
<td>spa-ESP (Spanish EU)</td>
<td>spa-MX (Span-Mexico)</td>
<td>spa-USA (Span-US)</td>
<td></td>
</tr>
<tr>
<td>Czech</td>
<td>cs</td>
<td>ces-CZE (Czech)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Danish</td>
<td>dk</td>
<td>dan-DNK (Danish)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Finnish</td>
<td>fi</td>
<td>fin-FIN (Finnish)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>French (Canadian)</td>
<td>fr-CA</td>
<td>fra-CAN (French (CAN))</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hebrew</td>
<td>he</td>
<td>NONE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hungarian</td>
<td>hu</td>
<td>hun-HUN (Hungarian)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Norwegian</td>
<td>no</td>
<td>nor-NOR (Norwegian)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polish</td>
<td>pl</td>
<td>pol-POL (Polish)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Russian</td>
<td>ru</td>
<td>rus-RUS (Russian)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swedish</td>
<td>sv</td>
<td>swe-SWE (Swedish)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turkish</td>
<td>tr</td>
<td>tur-TUR (Turkish)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Romanian</td>
<td>ro</td>
<td>ron-ROU (Romanian)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Security of customer data

- Speech vendor must match Oracle standards.
- Multiple levels of security reviews
- Offsite speech data causes risks.
- Special arrangements for timely parsing, extraction, and deletion
Engineering for the Future

• Center of Excellence
  – Developer guidelines
  – Competitive analyses
  – Social networking
  – Code samples
  – Collateral
  – Improvements

• Expansion into product families
• Expansion into application platforms
• Testing speech recognizer alternatives
Thank you!