Maltese on the Brink

Mike Rosner
Dept. Artificial Intelligence
University of Malta
mike.rosner@um.edu.mt
Acknowledgements

Ray Fabri, Inst. Linguistics, UoM
invited talk at LREC2010

and others

• Albert Gatt (Inst Linguistics)
• Mike Spagnol (Uni. Konstanz)
• Duncan Attard (ex. UoM ICS)
• Claudia Borg (UoM ICS)
• Alexandra Vella (UoM/Inst Linguistics)
Outline

• Maltese Language
• Status of Maltese Language
• Current Projects
• Issues
Maltese Language
Maltese Language

• Mixed Language
  – Arabic: kelb (dog)
  – Romance: karozza (car)
  – English: swiċċ; ners; owkej

• Latin script + some special characters
  – ċ, ġ, ĥ, ż, għ, ie

• Vowels are always written (unlike Arabic)
  – kiteb
Mixed Language

<table>
<thead>
<tr>
<th>SEMITIC</th>
<th>NON SEMITIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>root-based</td>
<td>stem-based</td>
</tr>
<tr>
<td>non-concatenative</td>
<td>concatenative</td>
</tr>
<tr>
<td>Arabic</td>
<td>Sicilian, Italian, English</td>
</tr>
</tbody>
</table>
Mixed Language

**SEMITIC**
Arabic Origin

l-ilsien Malti
df-tongue Maltese
the Maltese language

*more formal*

**NON SEMITIC**
Italian Origin

il-lingwa Maltija
df-tongue Maltese
the Maltese language

*less formal*
## SEMITIC ASPECT

\(l\ -\ s\ -\ n\) radicals (root consonants)

interdigation of vowels

<table>
<thead>
<tr>
<th>Form</th>
<th>Root Structure</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td><code>lisn</code> CVCVC</td>
<td>talk</td>
</tr>
<tr>
<td>II</td>
<td><code>lissen</code> CVCCVC</td>
<td>utter/say</td>
</tr>
<tr>
<td>V</td>
<td><code>tlissen</code> t-CVCCVC</td>
<td>be uttered</td>
</tr>
<tr>
<td></td>
<td><code>(i)lsien</code> CCVVC</td>
<td>tongue</td>
</tr>
<tr>
<td></td>
<td><code>(i)lsn-a</code> CCC-a</td>
<td>languages</td>
</tr>
<tr>
<td></td>
<td><code>tlissin-a</code> t-CVCCV-a</td>
<td>utterance</td>
</tr>
<tr>
<td></td>
<td><code>tlissin</code> t-CVCCVC</td>
<td>uttering</td>
</tr>
<tr>
<td></td>
<td><code>lissien</code> CVCCVVC</td>
<td>utterer</td>
</tr>
<tr>
<td></td>
<td><code>milsen</code> m-VCCVC</td>
<td>dictionary (Aquilina)</td>
</tr>
</tbody>
</table>

June 2010

Translingual Europe: Maltese
NON-SEMITIC
Stem lingw

<table>
<thead>
<tr>
<th>form</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>lingw-a</td>
<td>tongue-fsg</td>
</tr>
<tr>
<td>lingw-i</td>
<td>tongue-pl</td>
</tr>
<tr>
<td>lingw-a-ġġ</td>
<td>tongue-fsg-nom</td>
</tr>
<tr>
<td>lingw-ist-a</td>
<td>tongue-d-fsg</td>
</tr>
<tr>
<td>lingw-ist-i</td>
<td>tongue-d-pl</td>
</tr>
<tr>
<td>lingw-ist-ik-a</td>
<td>tongue-d-d-fsg</td>
</tr>
<tr>
<td>lingw-ist-ik-u</td>
<td>tongue-d-d-msg</td>
</tr>
<tr>
<td>lingw-ist-iċ-i</td>
<td>tongue-d-d-pl</td>
</tr>
<tr>
<td>bi-lingw-i</td>
<td>d-tongue-d</td>
</tr>
<tr>
<td>mono-lingw-i</td>
<td>d-tongue-d</td>
</tr>
</tbody>
</table>

‘tongue/language’
‘tongues’
‘parlance/diction’
‘linguist’
‘linguists’
‘linguistics’
‘linguistic’
‘linguistic(pl)’
‘bilingual’
‘monolingual’
Plural Formation

<table>
<thead>
<tr>
<th>Sound Plural</th>
<th>Broken Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>formed by suffixes:</td>
<td>change of stem drop of vowel</td>
</tr>
<tr>
<td>(a) Romance</td>
<td>qamar/qmura</td>
</tr>
<tr>
<td>karozza/karozzi (car)</td>
<td>tifel/tfal</td>
</tr>
<tr>
<td>tappit/tappiti (carpet)</td>
<td>ġdid/ġodda (new)</td>
</tr>
<tr>
<td>(b) Semitic</td>
<td>tappit/twapet (carpet)</td>
</tr>
<tr>
<td>ikla/ikliet (food)</td>
<td></td>
</tr>
</tbody>
</table>
Morpho-Syntactic Features

• Verbless sentences
  Il-karozza ġdid/the car is new
  cf. Il-karozza il-ġdid/the new car

• Construct state (inalienable possession)
  Id it-tifel/the boy's hand
  Kelb il-baħar/dog the sea = shark

• Sun-letters
  ix-xemx/the sun
  it-tifel/the boy
Clitic Pronouns

• bgħatthielux

• bgħat – t – hie – lu – x
• send past to her it not
  1SM

• I didn't send it to her
<table>
<thead>
<tr>
<th>Italian Borrowing</th>
<th>English Borrowing</th>
</tr>
</thead>
<tbody>
<tr>
<td>spjega explain (lt. spiegare)</td>
<td>ixxuttja kick a football (Eng. shoot)</td>
</tr>
<tr>
<td>jispjega he explains</td>
<td>jixxuttja he kicks</td>
</tr>
<tr>
<td>nispjegaw we explain</td>
<td>nixxuttjaw we kick</td>
</tr>
<tr>
<td>spjegat she explained</td>
<td>ixxuttjat she kicked</td>
</tr>
<tr>
<td>spjegajt I explained, etc.</td>
<td>ixxuttjajt I kicked, etc.</td>
</tr>
</tbody>
</table>
Language Situation in Malta
Status of the Maltese Language

According to the Constitution

• Maltese is the national language
• English and Maltese are official languages, i.e.
  – laws in EN and MT
  – all official government correspondence should be bilingual
• In reality there is a strong tendency for English to be used in government departments
• Maltese tends to be the spoken language
• English tends to be the written language
Language Spoken at Home
(2005 Census)
The Written Language
Since 2004

- Official EU Language
- Higher socio-political status
- Increased language awareness
- Need for
  - translation, translators, interpreters
  - developments in terminology
  - further standardisation of orthography
- Formation of National Council for the Maltese Language in 2005
Functions of Kunsill include

- Promotion the Maltese Language
- Updating orthography
- Adoption of a suitable linguistic policy
- Evaluation and coordination of the work done by associations and individuals in the Maltese language sector
- Establishment of a National Centre of the Maltese Language which, besides serving as the office of the Council, shall offer the necessary printed and audiovisual resources

Current Projects
Research, Resources and Tools

• Maltese Language Resource Server (UM)
• Maltobi and SPaN projects (UM)
• PsyCol Maltese Lexical Corpus (UM/U. Ariz.)
• Online lexicon/dictionary (UM/U. Ariz.)
• Text to Speech Engine (FITA/UM/CW)
• Spellchecker (UM/MS)
• CLARIN (UoM/FP7)
## Projects

<table>
<thead>
<tr>
<th></th>
<th>Lexicon</th>
<th>Corpus</th>
<th>Tools</th>
<th>Text</th>
<th>Speech</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLRS</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>MalTOBi</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>PsyCOL</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>TTS</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Spellcheck</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>CLARIN</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
Future Work: Basic Infrastructures

• Resources
  – Corpora: MNC + others
  – Lexical: mono and multi lingual; MaltiWordNet

• Tools
  – Tagger (improve accuracy)
  – Spellchecker
  – Morphological Analyser
  – Parser
  – Translation Support
Conclusions

• Mixed language hence more complex.
• Still lacking basic language resources.
• Lack of stable funding. Chicken and egg situation regarding users and funders of language-enabled tools and resources.
• Order in which to tackle shortcomings. Roadmap needed.
• Ensuring survival of Maltese in the midst of new technological developments.
• Balancing act between conservation and modernisation.
Valletta in Winter