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, ħ, 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>

June 2010 Translingual Europe: Maltese
## Mixed Language

<table>
<thead>
<tr>
<th>SEMITIC</th>
<th>NON SEMITIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabic Origin</td>
<td>Italian Origin</td>
</tr>
<tr>
<td>I-ilsien Malti</td>
<td>il-lingwa Maltija</td>
</tr>
<tr>
<td>df-tongue Maltese</td>
<td>df-tongue Maltese</td>
</tr>
<tr>
<td>the Maltese language</td>
<td>the Maltese language</td>
</tr>
<tr>
<td><em>more formal</em></td>
<td><em>less formal</em></td>
</tr>
</tbody>
</table>

June 2010 Translingual Europe: Maltese
SEMITIC ASPECT

*l - s - n* radicals (root consonants)
interdigitation of vowels

<table>
<thead>
<tr>
<th>Word</th>
<th>Root</th>
<th>Meaning</th>
<th>Form</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>lisen</em></td>
<td>CVVCV</td>
<td>‘talk’</td>
<td>I form</td>
</tr>
<tr>
<td><em>lissen</em></td>
<td>CVCCVC</td>
<td>‘utter/say’</td>
<td>II form</td>
</tr>
<tr>
<td><em>tlissen</em></td>
<td>t-CVCCVC</td>
<td>‘be uttered’</td>
<td>V form</td>
</tr>
<tr>
<td><em>(i)lsien</em></td>
<td>CCVVC</td>
<td>‘tongue’</td>
<td></td>
</tr>
<tr>
<td><em>(i)lsn-a</em></td>
<td>CCC-a</td>
<td>‘languages’</td>
<td></td>
</tr>
<tr>
<td><em>tlissin-a</em></td>
<td>t-CVCCV-a</td>
<td>‘utterance’</td>
<td></td>
</tr>
<tr>
<td><em>tlissin</em></td>
<td>t-CVCCVC</td>
<td>‘uttering’</td>
<td></td>
</tr>
<tr>
<td><em>lissien</em></td>
<td>CVCCVVC</td>
<td>‘utterer’</td>
<td></td>
</tr>
<tr>
<td><em>milsen</em></td>
<td>m-VCCVC</td>
<td>‘dictionary’</td>
<td>(Aquilina)</td>
</tr>
</tbody>
</table>
**NON-SEMITIC**

Stem lingw

<table>
<thead>
<tr>
<th>Stem</th>
<th>Example</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>lingw-a</td>
<td>tongue-fsg</td>
<td>‘tongue/language’</td>
</tr>
<tr>
<td>lingw-i</td>
<td>tongue-pl</td>
<td>‘tongues’</td>
</tr>
<tr>
<td>lingw-a-ġġ</td>
<td>tongue-fsg-nom</td>
<td>‘parlance/diction’</td>
</tr>
<tr>
<td>lingw-ist-a</td>
<td>tongue-d-fsg</td>
<td>‘linguist’</td>
</tr>
<tr>
<td>lingw-ist-i</td>
<td>tongue-d-pl</td>
<td>‘linguists’</td>
</tr>
<tr>
<td>lingw-ist-ik-a</td>
<td>tongue-d-d-fsg</td>
<td>‘linguistics’</td>
</tr>
<tr>
<td>lingw-ist-ik-u</td>
<td>tongue-d-d-msg</td>
<td>‘linguistic’</td>
</tr>
<tr>
<td>lingw-ist-iċ-i</td>
<td>tongue-d-d-pl</td>
<td>‘linguistic(pl)’</td>
</tr>
<tr>
<td>bi-lingw-i</td>
<td>d-tongue-d</td>
<td>‘bilingual’</td>
</tr>
<tr>
<td>mono-lingw-i</td>
<td>d-tongue-d</td>
<td>‘monolingual’</td>
</tr>
</tbody>
</table>
Plural Formation

Sound Plural
formed by suffixes:

(a) Romance
karozza/karozzi (car)
tappit/tappiti (carpet)

(b) Semitic
ikla/ikliet (food)

Broken Plural
change of stem
drop of vowel

qamar/qmura
tifel/tfal
gdid/ġodda (new)
tappit/twapet (carpet)
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
## Verbs with Semitic Inflections

### Italian Borrowing
- spjega explain (It. spiegare)
- jispjega he explains
- nispjegaw we explain
- spjegat she explained
- spjegajt I explained, etc.

### English Borrowing
- ixxuttja kick a football (Eng. shoot)
- jixxuttja he kicks
- nixxuttjaw we kick
- ixxuttjat she kicked
- ixxuttjajt I kicked, etc.
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)

- Total
- Maltese
- English
- Other
- >1 Language

June 2010

Translingual Europe: Maltese
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>X</td>
<td></td>
<td></td>
<td></td>
<td></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