

II-MALTI F'SALIB IT-TOROQ TAL- IZVILUPPI FIT-TEKNOLOĠIJA

MALTESE AT THE CROSSROADS OF TECHNOLOGICAL DEVELOPMENTS

Ray Fabri

L-Istitut tal-Lingwistika
L-Università ta' Malta

Institute of Linguistics
University of Malta

Irec Malta - Il-Ħamis, 20 ta' Mejju 2010

Irec Malta - Thursday, 20th May 2010

Merħba bikom

Merħba bi-kom
welcome with-you(pl)
(You are) welcome

- to the conference
- to Malta

Nirringrazzjakom

N-irringrazzja-kom

I -thank -you(pl)

I thank you

- all of you

- the organisers

il-Malti
the-Maltese

Maltese

- Senior lecturer in
General Linguistics
Lingwistika Ġenerali
- Chairman of the
Institute of Linguistics/L-Istitut tal-Lingwistika
University of Malta/L-Università ta' Malta
- Member of the National Council for the Maltese
Language
Il-Kunsill Nazzjonali tal-Ilsien Malti
- Head of Committee on Maltese and ICT
(within the Council)
Il-Kumitat Tekniku għall-Iżvilupp tal-Malti fl-
Informatika

What is peculiar to Maltese that makes it an interesting object of research in linguistics?

Semitic element
(mainly Arabic)
+
non-Semitic element
(mainly Sicilian, Italian, English)

non-concatenative, root-based
+
concatenative, stem-based
properties

l-ilsien Malti
df-tongue Maltese
the Maltese language

il-lingwa Maltija
df-tongue Maltese
the Maltese language

Arabic origin: root-based

l-ilsien Malti
df-tongue Maltese
the Maltese language

l - s - n radicals (root)

<i>lisen</i>	CVCVC	‘talk’	(1st form) (theme/conjugation/binyan)
<i>lissen</i>	CVCCVC	‘utter/say’	(2nd form)
<i>tlissen</i>	t-CVCCVC	‘be uttered’	(5th form)
<i>(i)lsien</i>	CCVVC	‘tongue/language’	
<i>(i)lsn-a</i>	CCC-a	‘tongues/languages’	
<i>tlissin-a</i>	t-CVCCV-a	‘utterance’	
<i>tlissin</i>	t-CVCCVC	‘uttering’	
<i>lissien</i>	CVCCVVC	‘utterer’	
<i>milsen</i>	m-VCCVC	‘dictionary’	(Aquilina)

NB: VV = long vowel, digraph ‘ie’ = [:]

Italian origin: stem-based
lingw-

il-lingwa Maltija
df-tongue Maltese
the Maltese language

lingw- stem

<i>lingw-a</i>	tongue-fsg	‘language’
<i>lingw-i</i>	tongue-pl	‘languages’
<i>lingw-a-gg</i>	tongue-fsg-nom	‘parlance/diction’
<i>lingw-ist-a</i>	tongue-der-fsg	‘linguist’
<i>lingw-ist-i</i>	tongue-der-pl	‘linguists’
<i>lingw-ist-ik-a</i>	tongue-der-der-fsg	‘linguistics’
<i>lingw-ist-ik-u</i>	tongue-der-der-msg	‘linguistic’
<i>lingw-ist-ic-i</i>	tongue-der-der-pl	‘linguistic(pl)’
<i>bi-lingw-i</i>	der-tongue-der	‘bilingual’
<i>mono-lingw-i</i>	der-tongue-der	‘monolingual’

N.B. der = derivational affix

l-ilsien Malti

generally more formal register

il-lingwa Maltija

generally less formal register

another example

għ-l-m

għallem 'teach'

għalliem 'male teacher'

għalliema 'female teacher'

għalliema 'teachers'

is-ser 'male teacher'

il-miss 'female teacher'

it-tiċer 'fe/male teacher'

Extract from the newspaper *It-Torça* (30/4/2006)

Noel Turner kien il-muttur tal-**Blues** f'*nofs* il-**gawnd** fejn wettaq ħafna *xogħol utli*, imma Valletta deheru **jikkontrollaw** tajjeb, minkejja li bdew *mingħajr* Sullivan u Fenech, li lanqas biss kienu fost is-**sostituti**... l-ewwel taqsima ... ntemmet mingħajr **gowls**. (<http://www.it-torca.com/>)

Noel Turner was the Blues' motor in midfield, where he did some very useful work; but Valletta seemed to control the game well, although they started without Sullivan and Fenech, who were not even present among the substitute players... The first half ended without any goals.

Vocabulary

Maltese word	English translation	Origin
utli	: useful	Italian : utile/i
jikkontrollaw	: they control	Italian : controllare
sostitut	: substitute	Italian : sostituto/i
Blues	: the Blues	English : Blues
gawnd	: ground/pitch	English : ground
gowls	: goals	English : goal/s
nofs	: middle	Arabic : nuf (half)
xogħol	: work	Arabic : xuġl (work)
mingħajr	: without	Arabic : min (from) + ġajr (exception)

The status of the Maltese language

The Constitution

Section 5 [National Language], chapter I:

the ‘national language of Malta is the Maltese language’

with ‘the Maltese and the English languages’ as ‘the official languages of Malta’

For example:

- laws are published both in Maltese and English
- in principle all 'official' (government etc.) correspondence and information should be bilingual
- in reality strong tendency for English to be used, in particular within government departments

In practice

- Maltese still tends to be the spoken language
- English still tends to be the dominant written language



OPEN

CLOSED

BACK IN 5 MINUTES

MIFTUH

MAGHLUQ

LURA 5 MINUTI OHRA

Recently:

two events that had a positive effects on

- the language itself
- people's perceptions/awareness



Since 2004

official language of the European Union

- higher socio-political status
- increased language awareness locally
- interest in language abroad
- rapid increase in number of professional translators

- lack of interpreters
- need for development of terminology
- need refinements in standardisation
(e.g. especially orthography)

Since 2005

National Council for the Maltese Language



IL-KUNSILL NAZZJONALI TAL-ILSIEN MALTI

<http://www.kunsilltalmalti.gov.mt/>

‘to promote the National Language of Malta and to provide the necessary means to achieve this aim’
(ACT No. V of 2004)

Principles

1. Maltese is the language of Malta and a fundamental element of the national identity of the Maltese people.

...



Functions of Council

1. promote the Maltese Language...
2. update the orthography of the Maltese Language...
- ...
5. adopt a suitable linguistic policy...
- ...
7. evaluate and co-ordinate the work done by associations and individuals in the Maltese language sector...
- ...
- 11. establish a National Centre of the Maltese Language** which, besides serving as the office of the Council, shall offer the necessary printed and audiovisual resources...



Unfortunately

- no proper adequate premises yet
- no full-time paid staff yet (volunteer basis)

The language situation in Malta

2005 National Census

Language spoken at home

Population over 10 years of age

Total
362,376
100%

Maltese
326,703
90.5%

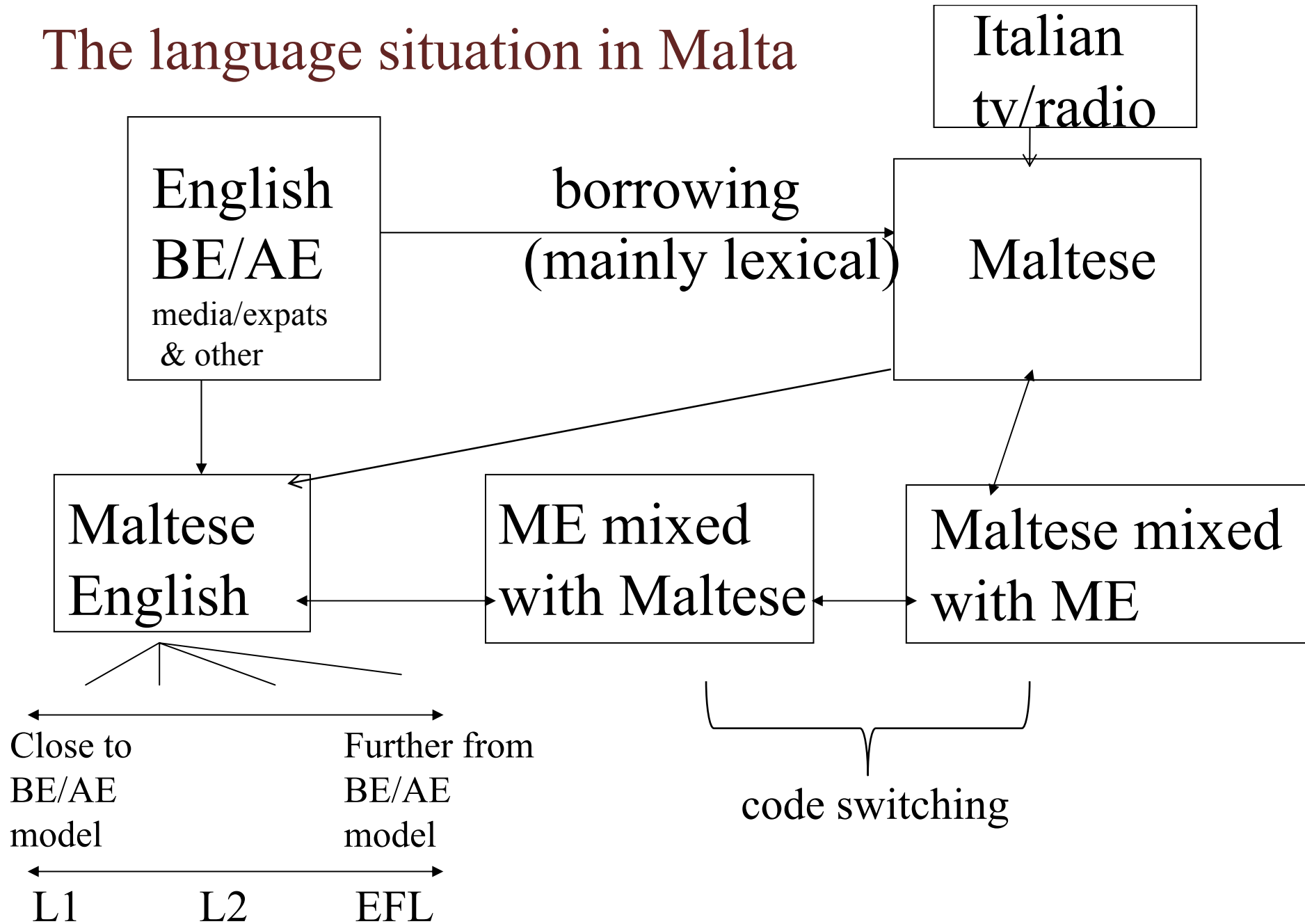
English
21,636
5.97%

other
3,344
0.92%

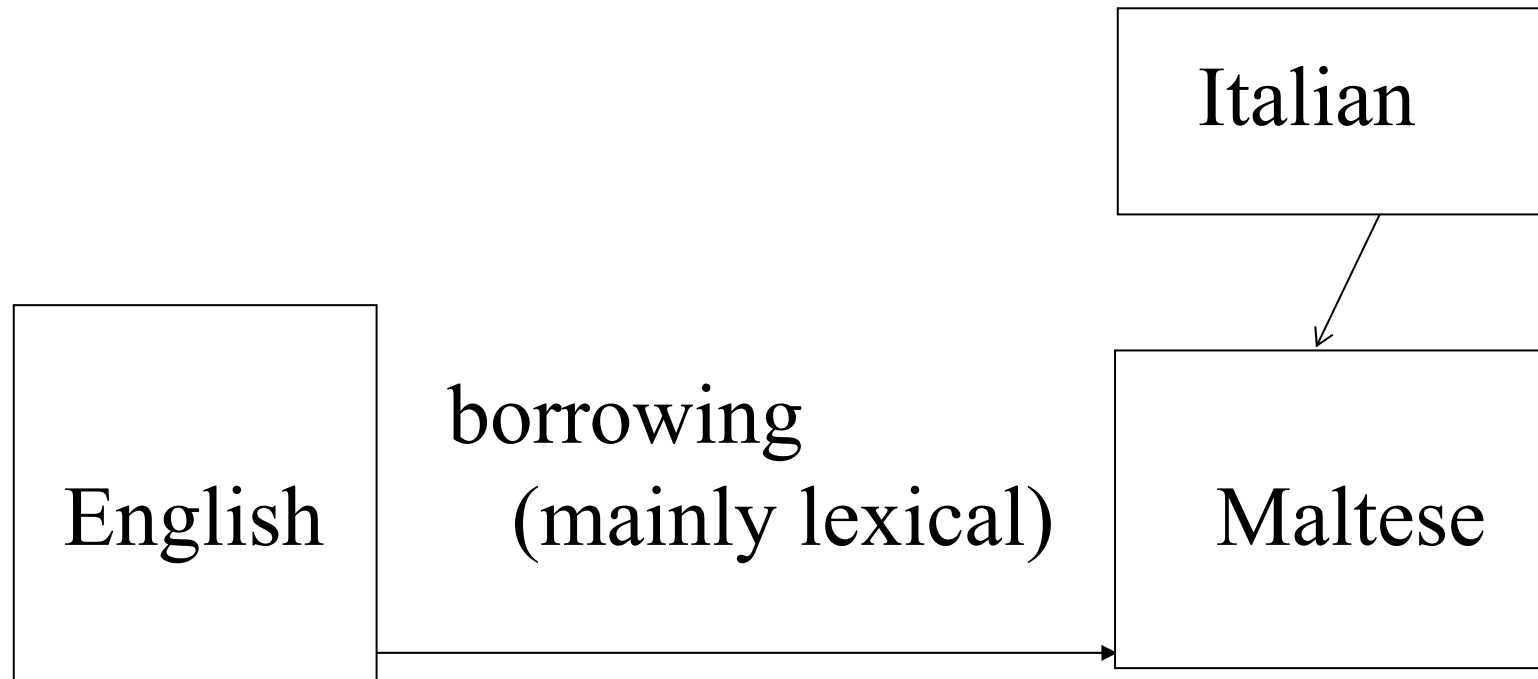
**more than
one language**
10,693
2.95%

Which language do you speak most at home?

The language situation in Malta



Borrowing and integration



Example of integration of noun from Italian:

broken plural

siġra
CVCCV

siġar
CVCVC

‘tree’ ‘trees’

villa - *vilel*

‘villa’ - ‘villas’

birra - *birer*

‘beer’ - ‘beers’

pizza - *pizez*

‘pizza’ - ‘pizzas’

Example of integration of verb from English *issejvja* ‘save (on computer)’

Model: *dara/jidra* ‘get used to’

I save/d	<i>nissejvja</i>	<i>issejvjajt</i>
you save/d	<i>tissejvja</i>	<i>issejvjajt</i>
he saves/d	<i>jissejvja</i>	<i>issejvja</i>
she saves/d	<i>tissejvja</i>	<i>issejvjat</i>
we save/d	<i>nissejvjaw</i>	<i>issejvjajna</i>
you save/d	<i>tissejvjaw</i>	<i>issejvjajtu</i>
they save/d	<i>jissejvjaw</i>	<i>issejvjaw</i>

Other examples: *jiskorja*, *jimmoniterja* ‘monitor’, *jiggugilja* ‘google’ etc.,

synthetic forms

ma nissejvjawhilux

ma n-issejvja-w-hi-lu-x

neg 1-save-pl-3fsg-3msg-neg

Sb Sb-dO -iO

We do not save it for him

synthetic forms and spell checking

ħarab

‘escape’

ħarrab

‘cause to escape’

n-ħarrab

‘I cause to escape’

n-ħarrb-u

‘we cause to escape’

n-ħarrb-u-ha

‘we cause her to escape’

n-ħarrb-u-hie-lha

‘we cause her to escape for her’

ma n-ħarrb-u-hi-lhie-x

‘we do not cause her to escape for her’

orthogaphy

especially words of English origin

e.g. futbol

BUT

hockey?? hoki??

airconditioner erkondixiner/erkondixin

- Maltese very dynamic, growing
- trick to balance conservation and loss with modernisation and enrichment
- make sure Maltese survives in the midst of new technological developments

Maltese at the crossroads

research, resources and tools

Corpora

3 projects

- 2 local

- 1 USA

1. MLRS (Maltese Language Resource Server)

- run by Dept of Intelligent Computer Systems and Institute of Linguistics (Mike Rosner, Albert Gatt, Ray Fabri)
- 'interacting' electronic lexicon & written corpus and basic language resources and tools
- corpus: ca. 9.5 million tokens
 - level of annotation:
 - Structural markup (paragraph, sentence, token)
 - Header indicating origin, author etc.
 - POS Tagging (currently at ca. 82% accuracy; tagger is being fine-tuned)
 - tools (search, tagging etc.)
 - planned to go public at end of this year 2010³⁹

2. MalToBI and SPaN projects

- speech annotation system for spoken corpora of Maltese, annotated at both orthographic and prosodic levels
- run Institute of Linguistics

Alexandra Vella

3. The PsyCoL Maltese Lexical Corpus (PMLC):

- approx. 3.5 million tokens
- University of Arizona (Adam Ussishkin)

<http://dingo.sbs.arizona.edu/~psycol/resources/>

J. Francom, A. Ussishkin, and D. Woudstra. 2009. Creating a Web-based Lexical Corpus and Information Extraction Tools for the Semitic Language Maltese. *Proceedings of the SALTMIL 2009 workshop on Information Retrieval and Information Extraction for Less Resourced Languages.*, pages 916.

On-line lexicons/dictionaries

1. MLRS (Maltese Language Resource Server)

- mainly with grammatical information
- still in its infancy

2. On-going work on-line Maltese English dictionary:

- a local company
- the National Council for the Maltese Language
- expertise from the University of Malta
- co-operation with a foreign university

Text-to-speech system

- private company after winning government tender
- own programmers + university expertise (software engineering and linguistics)
- still at the beginning

Iċ-ċekkjatur ‘spellchecker’

- Microsoft
 - to be launched be released within the Microsoft Office Language Interface Pack 2010
 - as yet no specific dates for the release of the Maltese Language Interface Pack 2010
- free on-line spell-checker
 - developed by Ramon Casha
 - available at <http://linux.org.mt/spellcheck>
(Malta Linux User Group))
- no sign yet of grammar and style checkers

Development of Infrastructure

The University of Malta participates in

CLARIN Project

- ‘a large-scale pan-European collaborative effort to create, coordinate and make language resources and technology available and readily usable.’

<http://www.clarin.eu/internal>

- aim: to create infrastructure which encourages the sharing and processing of such resources on a European scale in order to open new perspectives in Humanities and Social Sciences research
- Mike Rosner

Institute of Linguistics, University of Malta

- plan to introduce new course programme in **Human Language Technology**
- at BSc, MSc and PhD level
- parallel to General Linguistics programme
- include e.g. practical projects and work placements
- done our homework, waiting for answer from responsible units
- promise of new lecturing and research posts in area

Development of resources and tools moving at a slow pace

- lack of funds: no proper investment
- no obvious direct financial advantage
- lack of/limited local expertise
- prejudice
- ignorance: unaware of expertise, work and time required
- lip service

Maltese at the crossroads

choice has to be made NOW



1. along the road to further, speedy development - massive investment
2. down the road to neglect, foot-dragging - lack of investment

Grazzi ħafna