An online database of root-and-pattern verbs

John J. Camilleri — Chalmers / University of Gothenburg
Michael Spagnol — University of Malta

Verb formation strategies

Root-and-pattern

kiser
kisser
tkisser
nkiser

Concatenative

organizza
ssagrifika
pparkja
ppretenda

Root-and-pattern verbs

- ## Two intertwined morphemes
  - a consonantal root — 3 or 4 consonants
  - a verbal pattern (*binyan*)

Root	k-t-b	ħ-r-b-t
Pattern	1v2v3	1v23v4
Verb	*kiteb*	*ħarbat*
Gloss	"write"	"confuse"

Tri-consonantal root classes

- Strong
  - Regular (k-t-b *kiteb* "write")
  - Geminated (ħ-b-b *ħabb* "love")
- Weak
  - Initial (w-r-t *wiret* "inherit")
  - Medial (s-j-r *sar* "become")
  - Final (m-x-j *mexa* "walk")
- Irregular
  - ('-d-n *stieden* "invite")

Quadri-consonantal root classes

- Strong
  - (ħ-r-b-t *ħarbat* "confuse")
- Weak-final
  - (k-n-t-j *kanta* "sing")

Patterns / derived forms

- ## 9 tri-consonantal patterns

root	I	II	III	V	VI	VII	VIII	IX	X
d-n-b	*dineb*								
ħ-s-r		*ħassar*		*tħassar*					
f-r-x	*firex*	*ferrex*		*tferrex*		*nfirex*			

- ## 2 quad. patterns

root	I	II
ħ-r-b-t	*ħarbat*	*tħarbat*

Original source

Michael Spagnol. *A Tale of Two Morphologies: Verb structure and argument alternations in Maltese*. PhD thesis, University of Konstanz, 2011.

- Sourced from Maltese dictionaries
- Appendices
  - Exhaustive list of 1,923 roots and 4,142 verbs
  - Microsoft Excel (*.xls) format

strong					b	č	d	f	ġ	g	għ	h	ħ	j	k	l	**m**	n	p	q	r	s	t	v	w	x	ż	z

root	I	II	III	V	VI	VII	VIII	IX	X
m-ċ-ċ		meċċeċ							
m-d-d	medd	*medded		tmedd, *tmedded	tmedd, *tmedded	ntmedd	mtedd		
m-għ-d	magħad	*magħgħad	miegħed		tmiegħed	ntmagħad	mtagħad		
m-għ-k	*megħek, magħak		miegħek		tmiegħek				
m-għ-r	megħer		miegħer		tmiegħer				
m-għ-x			miegħex		tmiegħex				
m-għ-ż			miegħeż		tmiegħeż				
m-h-l			miehel		tmiehel				
m-ħ-ħ		maħħaħ							
m-ħ-r		meħħer							
m-ħ-t	maħat	maħħat		tmaħħat		ntmaħat			
m-k-ħ		mekkaħ		tmekkaħ					
m-k-k		mekkek		tmekkek					
m-l-ħ[1]		mellaħ		tmellaħ				mlieħ	
m-l-ħ[2]	moloħ								
m-l-l	mell			tmell					stmell, stmerr
m-l-q		mellaq							
m-l-s	miles	melles		tmelles				mlies	
m-l-t		mallat		tmallat					
m-n-għ	mana'					ntmana'			
m-n-s		mannas		tmannas					

Computational perspective

- The .xls format
- Natural for human users
- Not ideal for computational tasks
  - Includes formatting information
  - Mixes data and layout
  - Not easily searchable

- Need something more amenable to computational processing

Conversion to TSV

- Export from spreadsheet to tab-separated-values
- Hand-written scripts to clean up formatting

root	I	II	III	V	VI	VII
m-ċ-ċ		meċċeċ				
m-d-d	medd	*medded		tmedd, *tmedded	tmedd, *tmedded	ntmedd
m-għ-d	magħad	*magħgħad	miegħed		tmiegħed	ntmagħad
m-għ-k	*megħek, magħak		miegħek		tmiegħek	
m-għ-r	megħer		miegħer		tmiegħer	
m-għ-x			miegħex		tmiegħex	
m-għ-ż			miegħeż		tmiegħeż	
m-h-l			miehel		tmiehel	
m-ħ-ħ		maħħaħ				
m-ħ-r		meħħer				
m-ħ-t	maħat	maħħat		tmaħħat		ntmaħat

Next step: relational database

- Powerful search capabilities
- Optimized and fast
- Structural modelling (not just a flat file)

- Steps
  a. Design relational model (schema)
  b. Import from TSV to database
  c. Host database
  d. Provide search interface

Relational model

Root
- Radicals
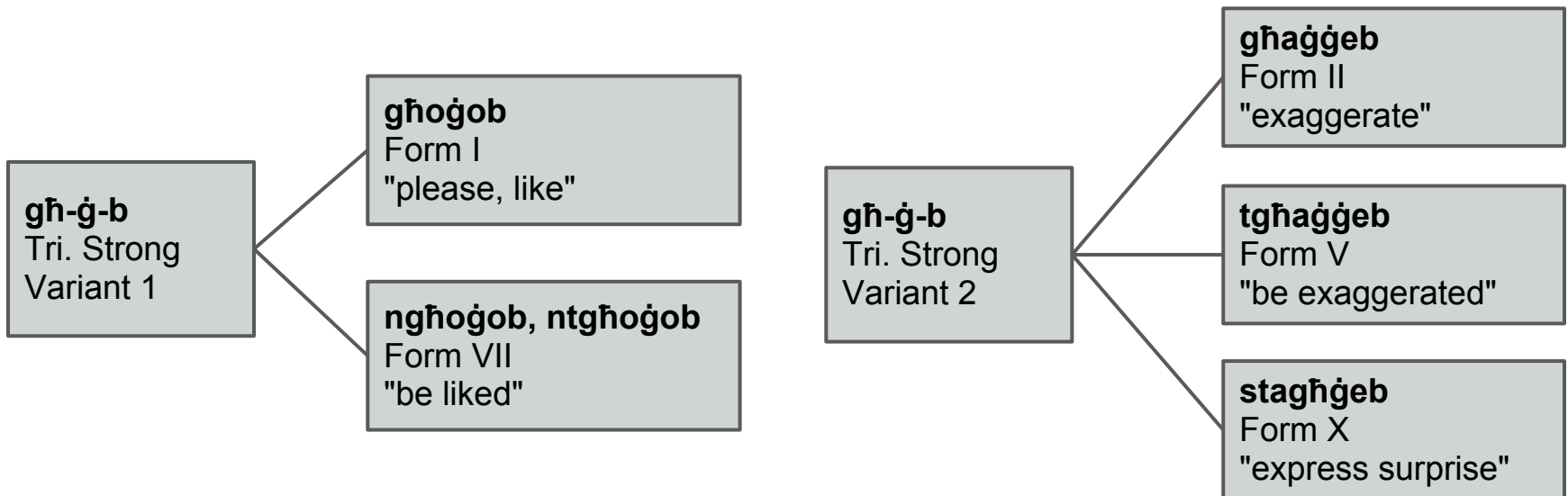- Alternate radicals
- Class
- Variant

1   n

Verb
- Surface form
- Alternate surface forms
- Derived form number
- Gloss
- Familiarity
- Transitivity (tr./intr./ditr.)
- Hypothetical

Relational model: examples

root	I	II	III	V	VI	VII	VIII	IX	X
għ-ġ-b[1]	għoġob					ngħoġob, ntgħoġob			
għ-ġ-b[2]		għaġġeb		tgħaġġeb					stagħġeb

Importing

- Roots and verbs
  - Automatic scripts from TSV
  - Manual post-editing
- Glosses
  - 1,046 entries in Excel spreadsheet
  - Convert .xls into TSV
  - For each, match root and verb form and insert gloss
  - Some errors corrected manually

The online database

- Entire verb list, accessible online
- Powerful search options
- Aimed at linguistics researchers & end users
- Hosted by the Maltese Language Resource Server (MLRS)

http://mlrs.research.um.edu.mt/resources/verbalroots/

Live demo

- Main page
  - Multilingual interface
- Basic browsing
- Verb detail page
- Searching
  - By free text
  - By root
  - By surface form
  - Advanced queries

A computational resource

- Web service
  - Interface with any other software application
- Offline use
  - Downloadable as MySQL database
  - For batch-processing (faster)
  - Transformable into other formats

The future

- Beginnings of an open computational lexicon
  - Easily imported into other systems
  - Welcome your own lists so that we can expand the database beyond just root-and-pattern verbs
- Combination with computational grammar
  - See my other talk!
- Supplementary data still very sparse
  - Contributors welcome

Acknowledgements

Thanks to Albert Gatt and Claudia Borg for hosting on MLRS.

...let's eat!