Effects of frequency and morphological structure - Investigating Maltese sound and broken plurals

Jessica Nieder & Holger Mitterer & Ruben van de Vijver

nieder@phil.hhu.de, holger.mitterer@um.edu.mt, Ruben.Vijver@hhu.de

Heinrich-Heine-Universität, Düsseldorf
DFG Research Unit FOR 2373: Project MALT

University of Malta, 21 November 2019

Morphology

Introduction

- study of internal word structure
- morphemes are smallest units: root morphemes vs. affixes
- there are different types of affixes (prefix, suffix, infix, etc.)

walk$_{\text{root}}$ -ed$_{\text{affix}}$
un-$_{\text{affix}}$ lock$_{\text{root}}$

Morphology

Concatenative vs. Nonconcatenative Morphology

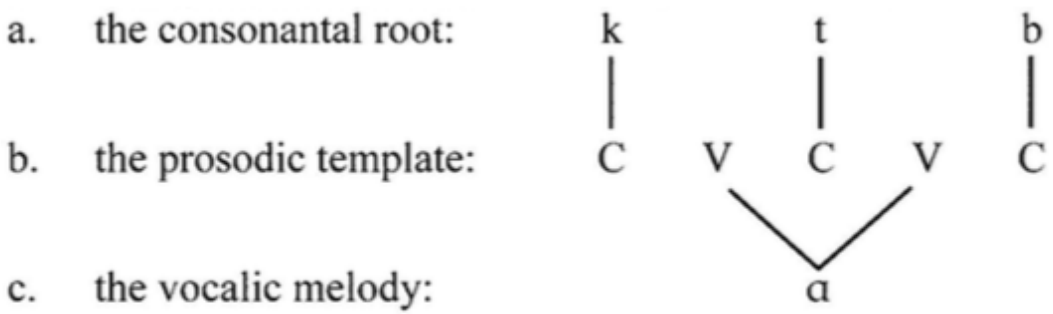
- concatenation: process of linear combination of morphemes to create new lexical items (e.g. affixation, compounding)
- English morphology can be described as predominantly concatenative: walk + -s, un- + do, house + work...
- but:
  swim $\rightarrow$ swam $\rightarrow$ swum
  goose $\rightarrow$ geese
- non-concatenative processes: internal changes/alternations of morphemes

- we find non-concatenative processes in Semitic languages like Arabic:

Arabic verb	Gloss
katab	he wrote
kutib	it was written
ʔaktub	he dictated
ʔuktab	it was dictated
kattab	he caused to write
kaatab	he corresponded

Table 1: Paradigm for the Arabic verb *katab* (Ussishkin, 2005)

- No string of segments - How to analyze the verb *katab*? - consonantal root as basis for analysis!
- possible representation (Ussishkin, 2005):



Maltese

- Maltese: An example of non-concatenative Morphology!

- 2 main strategies to build the plural of a noun:
  Sound Plural: concatenative via suffixation
  annimal – *annimali* 'animal(s)'
  Broken Plural: non-concatenative via internal restructuring of singular stem
  ballun – *blalen* 'ball(s)'

Maltese Plurals: The "problem"

Maltese speakers are faced with a dichotomy in their morphological system:

- sound plurals are built concatenatively by adding a suffix to the singular
- broken plurals are formed non-concatenatively by changing the prosody of the singular stem

Maltese Plurals: The "problem"

- overall we find more sound plurals than broken plurals in Maltese (based on the used data set)
- our data set: 2247 nouns in total
  - 1499 sound plurals = 67%
  - 758 broken plurals = 33%

Maltese Plurals: The "problem"

- How can we account for the choice of plural forms?

Production Experiment

Research Question

Is it possible to predict pluralisation of novel words? Can novel items be classified as broken or sound plurals?

- 3 steps: Data Set - Production Experiment (- NDL modeling)

Experimental Methods

One step back...

- Why do we need experiments in linguistics?
- What exactly is a linguistic experiment?
- What kind of experiments could we use?

Experimental Methods

What is an experiment?

- A procedure where we measure a specific effect we are interested in
  - based on variables we define
  - (usually) with a theoretical background in mind

Experimental Methods

Why do we need experiments?

- they help us to investigate the structure/processing/use of languages
- they help us to collect the data we need
- they help us to test linguistic theories

Experimental Methods

What kind of experiments could we use?

- online or offline methods
  - online = looking at language as it happens, e.g. eye-tracking, reaction time, perception or production experiments...
  - offline = looking at knowledge about language, e.g. questionnaires

Production Experiment

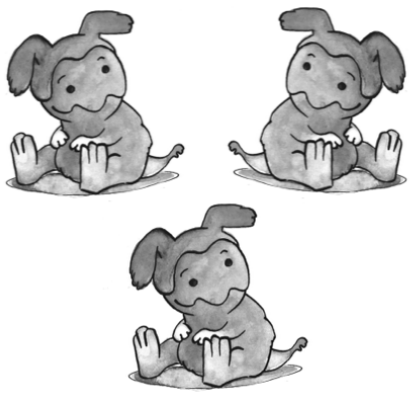
Method

- production task with visual presentation
- Maltese native speakers were asked to produce plural forms for existing Maltese singulars and phonotactically legal nonce singulars (Berko, 1958)
- nonce forms were constructed from words of our data set of 2373 Maltese nominals by changing either the consonants or the vowels or both systematically, e.g.: *sema* ‚sky', $\rightarrow$ *fera soma fora*
- the words used as base had either a sound plural form, a broken plural form or both plural forms

Production Experiment

Procedure

Dik l-istampa ta' snif



ħafna _____

Figure 1: Example for one trial of our experiment

Production Experiment

Stimuli

We chose **90 nonce words**:

- 30 from list C
  - 10 Base Broken Plural
  - 10 Base Sound Plural
  - 10 Base Both
- 30 from list V
  - 10 Base Broken Plural
  - 10 Base Sound Plural
  - 10 Base Both
- 30 from list CV
  - 10 Base Broken Plural
  - 10 Base Sound Plural
  - 10 Base Both

And **22 existing nouns**:

- 5 frequent sound plural words, 5 infrequent sound plural words
- 5 frequent broken plural words, 5 infrequent broken plural words
- 2 training items (1 sound plural, 1 broken plural)

Production Experiment

Results

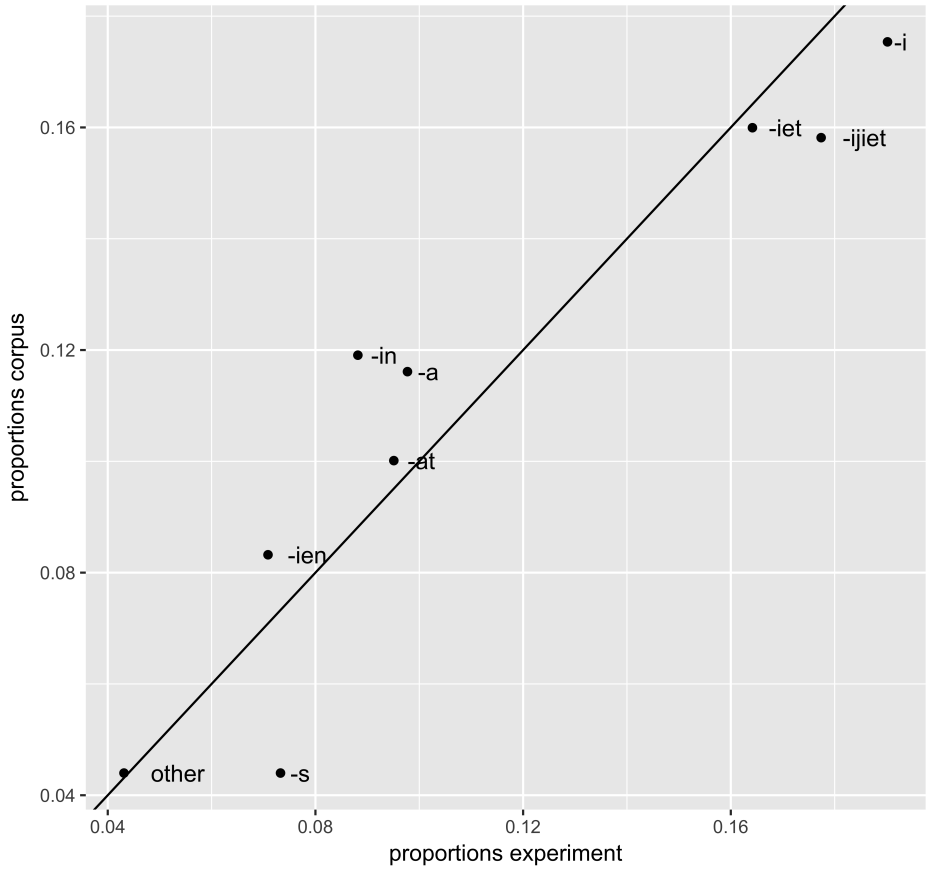


Figure 2: Proportion of different sound plural suffixes in corpus and experiment and their correlation.

regression analysis: multiple R-squared = 0.86, adjusted R-square = 0.84, F(45.72, 1), df = 7, p = .0002

Production Experiment

Results

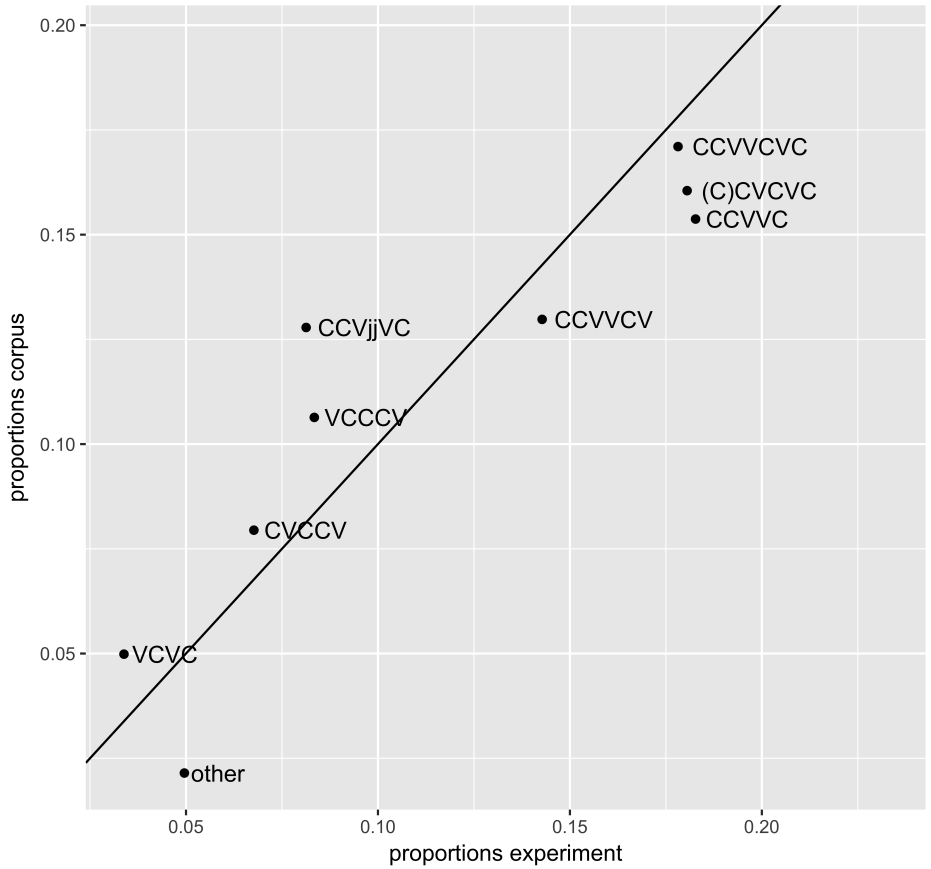


Figure 3: Proportion of different broken plural patterns in corpus and experiment and their correlation.

regression analysis: multiple R-squared = 0.82, adjusted R-square = 0.79, F(31.06, 1), df = 7, p = .0008

Maltese Experiment

Results - List

Does the change of consonants, vowels or both to build nonce words have an effect on the produced plural type of the nonce words?

Maltese Experiment

Results - List

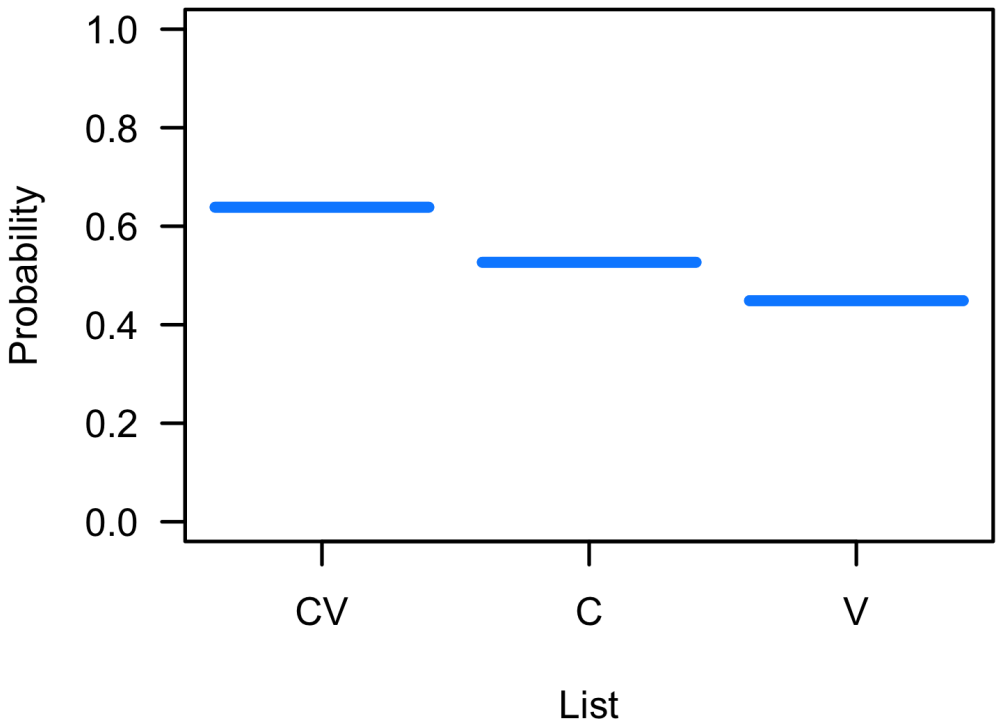


Figure 4: Results of a glmer model with variable: List

Significant difference between List CV and List V ($p<0.001$)

Maltese Experiment

Results - Base

Does the plural form of the existing word that has been used as a base for the nonce word have an effect on the produced plural type of the nonce words?

Maltese Experiment

Results - Base

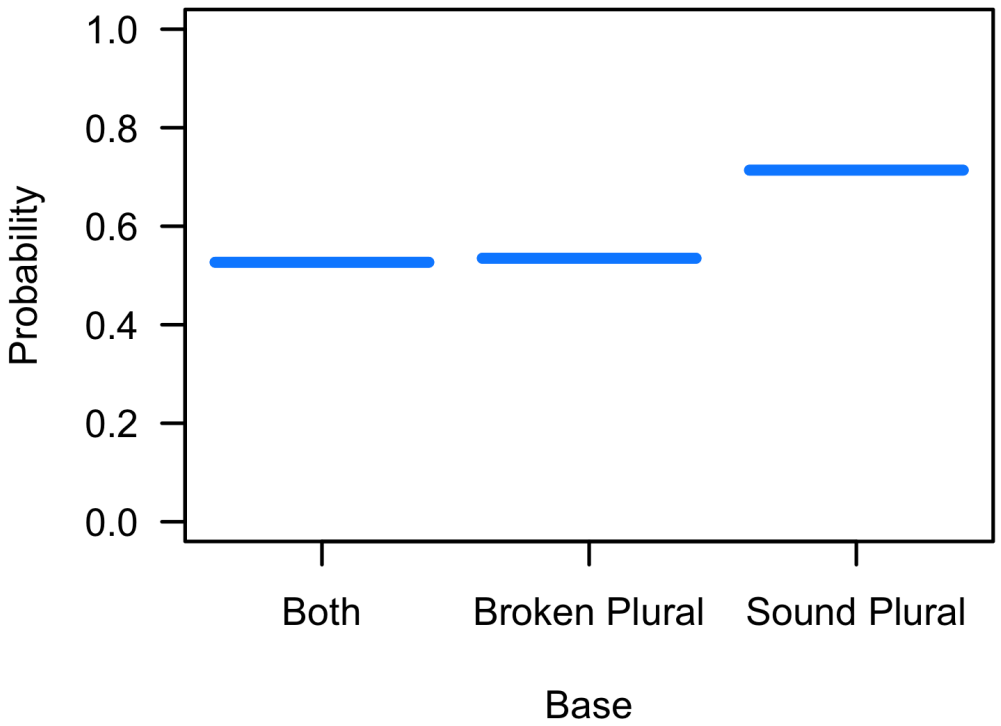


Figure 5: Results of glmer model with variable: Base

Significant difference between Base Broken and Base Sound ($p<0.001$)

Production Experiment

Results

- native speakers are able to generalize to novel nouns and use the most common suffixes and CV patterns for this task
- knowledge of Maltese native speakers of the singular-plural mappings is generalized to novel words on the basis of the similarity of the novel word to existing singulars and their associated plural form

Production Experiment

Results

- for nonce stems our participants used sound plurals more frequently:
  - 5404 sound plurals
  - 1262 broken plurals
- this is in line with the proportions we find in our data set

Reaction Time Experiment

Research Question

How are broken and sound plural forms represented in the Mental Lexicon?

- one possibility to test this: reaction time study with priming task

Dual-Mechanism Accounts

(e.g. Pinker, 1991, 1998; Pinker & Ullmann, 2002)

- two distinct mechanisms for processing complex word forms
- rule-based system: regular word forms
- whole word storage of irregular word forms

vs.

Single-Mechanism Accounts

(e.g. Rumelhart & McClelland, 1986; Skousen, 1992; Daelemans, 2002)

- irregular and regular complex word forms are processed within the same single mechanism

Dual Mechanism Accounts: Words and Rules Theory

(Pinker & Ullmann, 2002)

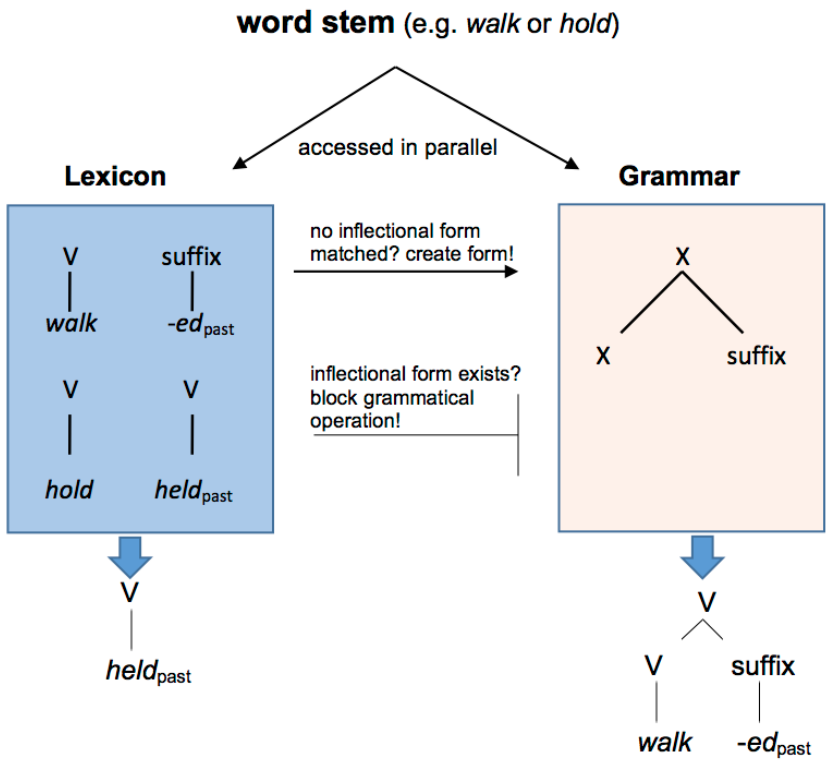
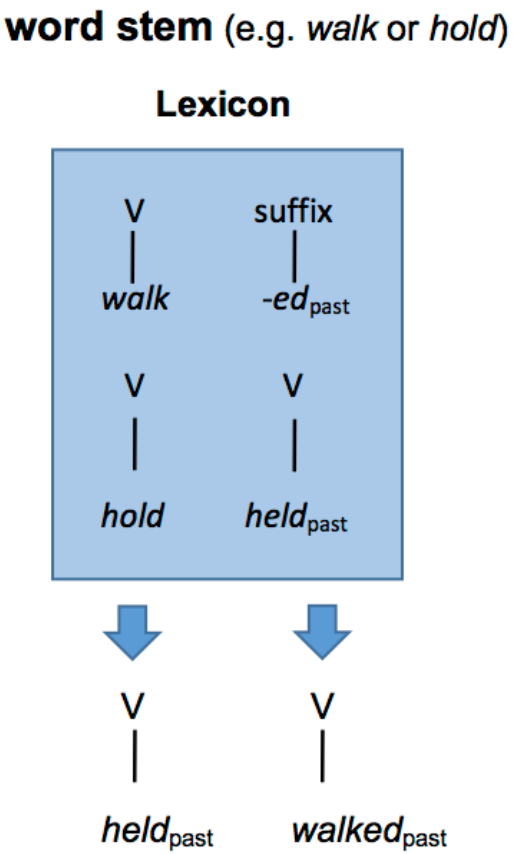


Figure 6: Words and Rules theory (adapted from Pinker & Ullmann, 2002)

Single Mechanism Accounts: Whole Word Storage



Reaction Time Experiment

Reaction time studies

- reaction time studies can be used to investigate spoken word processing
- RT (reaction time) is the time required to respond to a stimulus
- RT studies: experiments that involve the measurement of the amount of time participants need to respond to a stimulus
- RT is collected through simple tasks: lexical decision (simple yes/no questions: is *plimpa* a real word or not?)
- priming: experiments that involve the presentation of a related or unrelated prime (e.g. a word) before a target word
- different types of priming: intra-modal vs. cross-modal (see Justus, Yang, Larsen, de Mornay Davies & Swick, 2009, for an overview)

Reaction Time Experiment

Reaction time studies

- Sonnenstuhl, Eisenbeiss & Clahsen (1999) RT study on German: comparison of regular -s plurals, e.g. *Kino* - *Kinos*, to irregular -er plurals, e.g. *Mann* - *Männer*
  - they found differences in processing, data supports dual-mechanism account
- Meunier & Marslen-Wilson (2004) RT study on French regular and irregular verbs: *amener-amène* 'to bring - I bring' vs. irregular forms *peindre-peignent* 'to paint-they paint'
  - similar priming, data supports single-mechanism account
- Kielar, Joanisse & Hare (2008) RT study on English past tense forms
  - regular past tense forms and irregular past tense forms show consistent priming effects, data supports single-mechanism account

Maltese Reaction Time Experiment

Method

- reaction time experiment with cross-modal priming: auditory presentation of primes, visual presentation of targets
- 144 target items, two types of primes: Corresponding plural primes and phonologically and semantically unrelated control primes with the same plural suffixes or pattern like the corresponding plural word
- 2 frequent and 2 infrequent sound plural suffixes, 2 frequent and 2 infrequent broken plural patterns
- 144 nonce words as filler items
- 59 adult native speakers of Maltese (34 women and 25 men) performed a lexical decision task

Reaction Time Experiment

Stimuli

	PrimeType			
Target	Related Plural	Control Plural	Frequency	Plural Type
kappella	*kappelli*	*politiki*	high	sound
patri	*patrijiet*	*universitajiet*	high	sound
alla	*allat*	*triqat*	low	sound
qattiel	*qattiela*	*halliema*	low	sound
farfett	*friefet*	*xwabel*	high	broken
tifel	*tfal*	*swieq*	high	broken
storja	*stejjer*	*ktajjen*	low	broken
banda	*bnadi*	*ċrieki*	low	broken
vilnu	*vilel*	-	(filler)	(filler)

Table 2: Example of items that were used in the present reaction time study. Please note that the last row displays fillers (target = nonce words, prime = existing plural)

Reaction Time Experiment

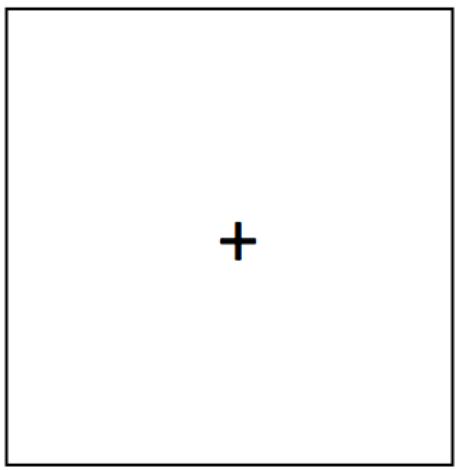
Predictions

Dual-Mechanism Hypothesis:
different frequency effect, difference in speed of processing

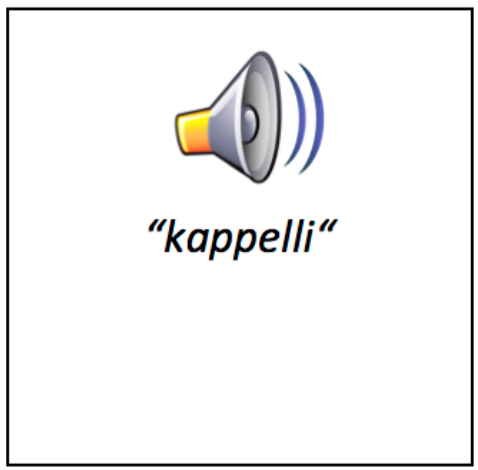
Single-Mechanism Hypothesis:
same frequency effect for both plural types, no difference in speed of processing

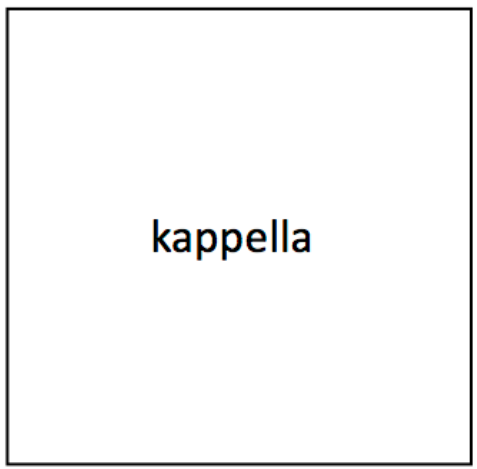
Reaction Time Experiment

Procedure



500ms





2000ms

Is it a real word? If yes, click the green button. If not, click the red button

Reaction Time Experiment

Results

Condition	Example	**Mean RT**
broken, plural prime	*qattus - qtates*	627
sound, plural prime	*omm - ommijiet*	630
broken, control prime	*ballun - fkieren*	670
sound, control prime	*vjaġġ - kuluri*	704
filler	*kapla - kapep*	776

Table 3: Mean reaction times

Reaction Time Experiment

Results

- we fitted several linear mixed effect regression models using the `lme4` package (Bates, Mächler, Bolker & Walker, 2015) in the R environment (R Core Team, 2016)
- dependent variable: log-transformed RT
- independent variables: PluralType, PrimeType, PluralFrequency, Origin, TargetFrequency, PrimeFrequency, Syllable
- random effects: Participant and Item
- PluralType, PrimeType, PluralFrequency were entered in the model as contrasts (Helmert coding)

Reaction Time Experiment

Results: First Model

	Estimate	Std. Err	t-value	p-value
Intercept	6.489808	0.024523	264.645	<2e-16 ***
PrimeFrequency	-0.031785	0.004920	-6.460	2.97e-10 ***
PluralType	0.027899	0.022605	1.234	0.218
PrimeFrequency:PluralType	0.001253	0.009460	0.132	0.895

Table 4: lmer model results: effect of the interaction of plural type and word frequency of plural forms on RT

Reaction Time Experiment

Results: First Model

- no significant interaction of PrimeFrequency and PluralType ($p = .9$) = no frequency effect
- our data supports a single-mechanism model of morphological processing
- question: What factors determined the observed reaction times for sound and broken plurals in Maltese?

Reaction Time Experiment

Results: Second Model

	Estimate	Std. Error	t-value	p-value
Intercept	6.524636	0.027391	238.202	<2e-16***
PrimeType	-0.094183	0.007289	-12.921	<2e-16***
PatternFrequency	-0.048664	0.016829	-2.892	0.004434 **
PluralType	0.022168	0.015801	1.403	0.162779
TargetFrequency	-0.033620	0.006122	-5.492	1.84e-07 ***
PrimeType:PatternFrequency	0.024779	0.011001	2.252	0.025949 *
PrimeType:PluralType	-0.046186	0.011426	-4.042	0.000135 ***

Table 5: Summary of the full model

Reaction Time Experiment

Results: Second Model

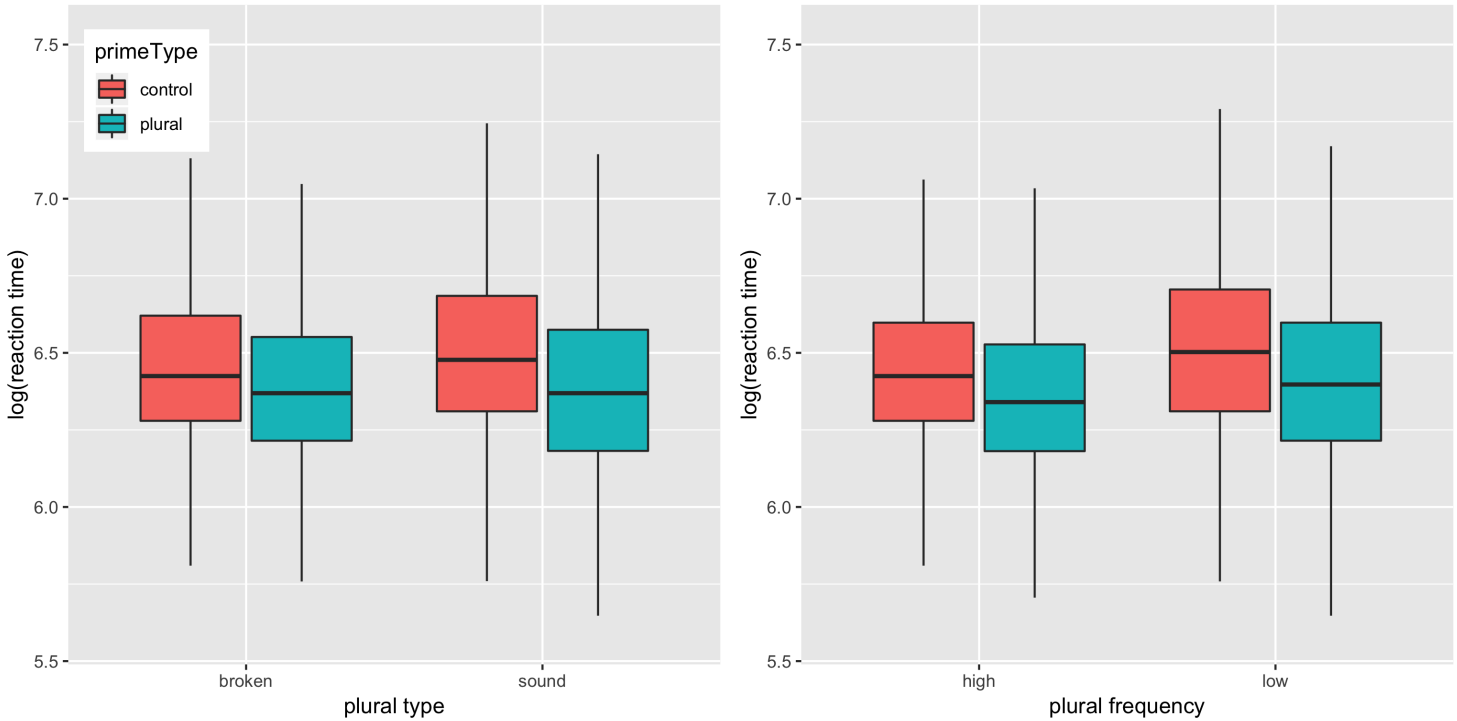


Figure 7: Effect of prime and plural type on RT (left); Effect of frequency of patterns and prime on RT (right)

- significant interaction between PrimeType and PluralType, significant interaction between PrimeType and PatternFrequency
- low frequency patterns elicited longer rt (Estimate: -0.049), decrease of rt the higher the frequency of the target words (Estimate: -0.034)

Reaction Time Experiment

Results: Second Model

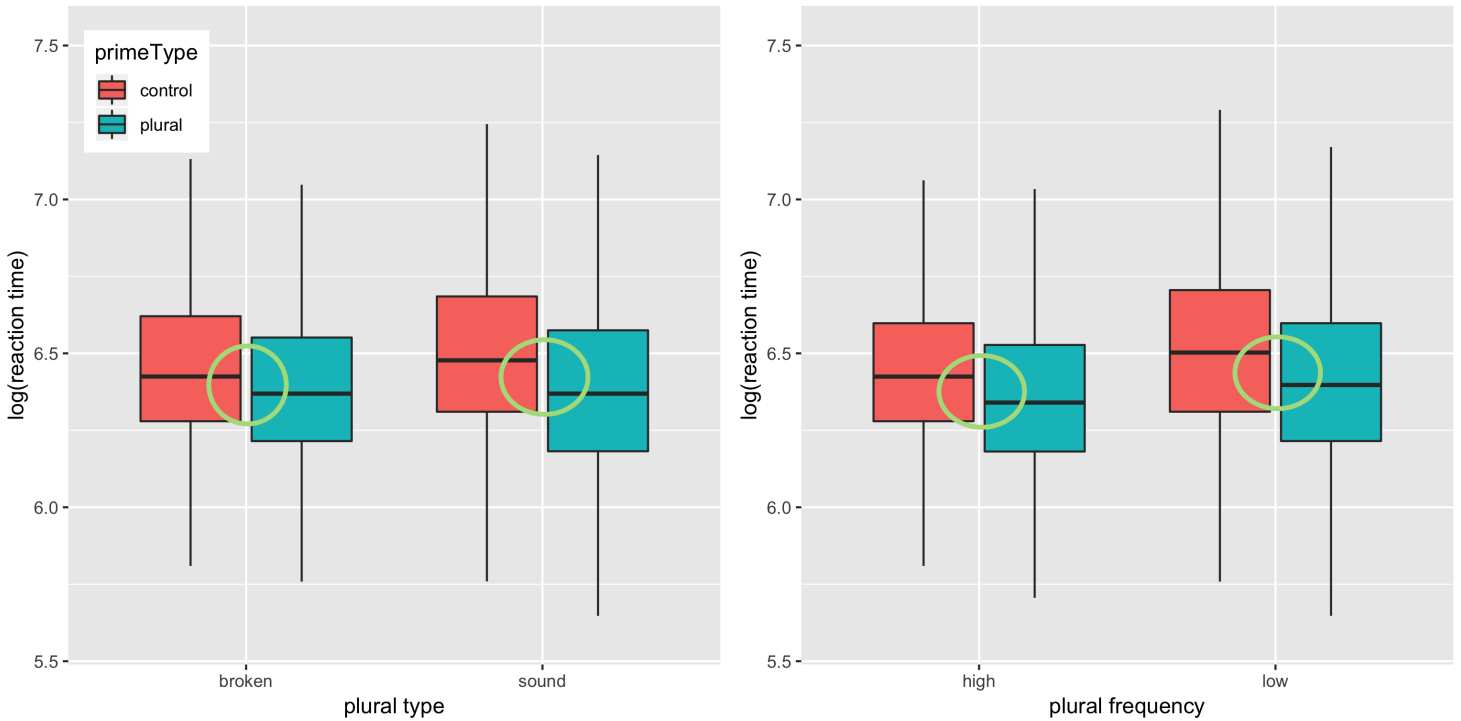


Figure 8: Effect of prime and plural type on RT (left); Effect of frequency of patterns and prime on RT (right)

- different priming effect

Reaction Time Experiment

Results: Second Model

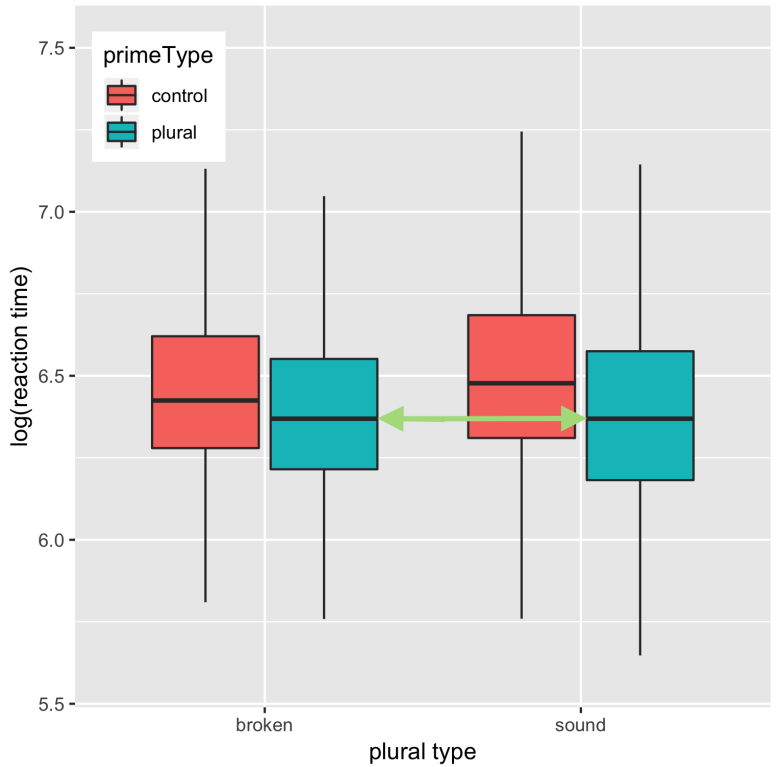


Figure 9: Effect of prime and plural type on RT

- no significant difference between broken and sound plurals

Reaction Time Experiment

Conclusion

- results indicate that Maltese broken and sound plurals are processed in the same way = single-mechanism account
- no difference in processing, no different word frequency effect; but: different priming effect!
  - greater priming for sound plurals due to phonological overlap of singular targets with their sound plural form
  - phonological overlap facilitates response latencies (Pastizzo & Feldman, 2002)
- instead of morphological regularity: frequency of patterns and the morphophonological similarity to related word forms are important factors for processing Maltese plurals

Discussion & Conclusion

The variation found in the nominal system influences the intuitions of native speakers about Maltese plural forms:

- Maltese native speakers are more certain about sound plural forms and use these forms more frequently.
- frequency effect: certain suffixes and patterns are more frequent than others and are thus faster to access in the mental lexicon.
- the nominal system of Maltese is split in two parts; generalizations concerning sound plurals are based on different information than generalizations for broken plurals.

Grazzi ħafna!

References I

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1).

Berko, J. (1958). The child's learning of english morphology. *WORD*, *14*(2), 150–177.

Daelemans, W. (2002). *A comparison of analogical modeling to memory-based language processing*, (pp. 157–179). Benjamins.

Drake, S. & Sharp, R. (2017). Productivity of the broken plural in maltese. In *6th International Conference of Maltese Linguistics*.

Farrugia, A. & Rosner, M. (2008). Maltimorph, a computational analysis of the maltese broken plural. *Proceedings of Workshop on ICT 2008*.

Justus, T., Yang, J., Larsen, J., de Mornay Davies, P., & Swick, D. (2009). An event-related potential study of cross-modal morphological and phonological priming. *J Neurolinguistics*, *22*(6), 584–604.

References II

Kielar, A., Joanisse, M. F., & Hare, M. L. (2008). Priming english past tense verbs: Rules or statistics? *Journal of Memory and Language*, *58*, 327–346.

Meunier, F. & Marslen-Wilson, W. (2004). Regularity and irregularity in french verbal inflection. *Language and Cognitive Processes*, *19*(4), 561–580.

Pastizzo, M. J. & Feldman, L. B. (2002). Discrepancies between orthographic and unrelated baselines in masked priming undermine a decompositional account of morphological faciliation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *28*(1), 244–249.

Pinker, S. (1991). Rules of language. *Science*, *253*(5019), 530–535.

Pinker, S. (1998). Words and rules. *Lingua*, *106*, 219–242.

Pinker, S. & Ullmann, M. T. (2002). The past and future of the past tense. *Trends in Cognitive Science*, *6*(11), 456–463.

References III

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rumelhart, D. E. & McClelland, J. (1986). On learning the past tenses of english verbs. In D. E. Rumelhart & J. McClelland (Eds.), *Parallel distributed processing*. MIT Press.

Skousen, R. (1992). *Analogy and Structure*. Springer Netherlands.

Sonnenstuhl, I., Eisenbeiss, S., & Clahsen, H. (1999). Morphological priming in the german mental lexicon. *Cognition*, *72*, 203–236.

Ussishkin, A. (2005). A fixed prosodic theory of nonconcatenative templatic morphology. *Natural Language and Linguistic Theory*, *23*, 169–218.

Acknowledgements

We thank Holger Mitterer for offering us the opportunity to use the Cognitive Science Lab at the University of Malta for conducting our experiment. We thank our colleagues from the DFG-Research Unit FOR2373 and our colleagues from the *Għaqda Internazzjonali tal-Lingwistika Maltija* for their advice and feedback.