Maltese Plurals: Evidence from a Nonce Word Experiment

Jessica Nieder & Ruben van de Vijver

nieder@phil.hhu.de, Ruben.Vijver@hhu.de

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

11th Mediterranean Morphology Meeting 2017
University of Cyprus, 22-25 June 2017
Maltese

- Semitic language with characteristics of Maghrebi Arabic, influenced by Sicilian, Italian and English
- National language of Malta, other official language: English
- Spoken by about 400,000 people
2 main strategies to build the plural of a noun:

- **Sound Plural**: concatenative via suffixation
  - *sptar* – *sptarijiet* 'hospital(s)'
- **Broken Plural**: non-concatenative via internal restructuring of singular stem
  - *ballun* – *blalen* ‘ball(s)’

There is variation within the two different plural forms:
- a number of sound plural suffixes, between 4 and 39 different broken plural patterns
- There is also variation in the choice of the plural forms:
  - *bandiera* (sg.) *bnadar* (broken pl.) vs. *bandieri* (sound pl.) ‘flag’
Maltese Plurals
Learnability

- Is it possible to predict pluralisation of novel words?
- If there are no rules governing the plural formation (Sutcliffe (1924) cited in Schembri (2012)), this means that there is no – linguistic or statistical – structure in the data that allows native speakers to generalize
Maltese Plurals
Previous accounts

Prosodic Morphology (McCarthy & Prince, 1996): Plural forms are mapped on prosodic templates (shape-invariant patterns)

- What happens in a system that shows a lot of variation?
- We find marked prosodic patterns: CCVV
- How to account for these patterns?

Dawdy-Hesterberg & Pierrehumbert (2014):
Ernestus & Baayen (2003) have shown that phonological features play a role for morphological generalization
**CV-skeleton mapping**

Has been used as description of different broken plural types in Maltese (e.g. Schembri (2012))

- How to account for sound plural forms?
- What skeletons trigger choice of plural forms?
Maltese Plurals

Previous accounts

- Common idea of these accounts: the phonotactics of the singular determines the shape of the (broken) plural
- This is a good starting point for both plural forms
1. The phonotactics of the singular determines the shape of the plural
2. More frequent items are more likely to be generalized than infrequent items
To test the hypotheses we created a corpus and conducted a production experiment.

We modeled our experimental data with the *Naive Discriminative Learner*, a cognitive learning algorithm (Baayen, Milin, Đurđević, Hendrix & Marelli, 2011) that does not rely on abstract representations like CV-structure: are generalizations possible?
We created a corpus of 2369 Maltese nominals.

Words were taken from Schembri (2012) and an online corpus (MLRS Corpus Malti v. 2.0).

Checked with Ġabra: online lexicon for Maltese (Camilleri, 2013).

CV structure.

Corpus frequency number for each word.
Figure 1: Distribution of Plural Types in our Corpus
Maltese 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 corpus of 2369 Maltese nominals by changing either the consonants or the vowels or both systematically, e.g.: *sema* 'sky', → *fera*
  *soma fora*
- The results are three lists of wug words: C, V, CV
- The words of our corpus used as base had either a sound plural form, a broken plural form or both plural forms: SP, BP, BOTH
Maltese 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)
Maltese Experiment
Procedure

- Participants: 80 adult native speakers of Maltese: 50 female, 30 male (mean age 24.6), recruited at the University of Malta
- We recorded the plural answers of the participants
- Steps: training phase, instructions in Maltese, test phase
- Stimuli were presented in randomized order
Dik l-istampa ta’ telleb
Ħafna _______________
There is a lot of variation in our data: different plural forms per item (broken plural = red, sound plural = green)
Does the change of consonants, vowels or both to build nonce words have an effect on the produced plural type of the nonce words?
Figure 2: Distribution of Plural Types within the lists C, CV and V
Maltese Experiment
Results - List

glmer with lme4 package (Bates, Mächler, Bolker & Walker, 2015)

- dependent variable:
  Answers of participants (binary, Sound or Broken Plural)

- independent variables:
  List = C, V, CV
  Base = SP, BP, BOTH

- random effects:
  Singular, Speaker
Maltese Experiment

Results - List

Figure 3: Results of glmer model with variable: List

Significant difference between List CV and List V (p<0.001)
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?
Figure 4: Distribution of Plural Types - Base
Figure 5: Results of glmer model with variable: Base

Significant difference between Base Broken and Base Sound (p<0.001)
Maltese Experiment
Results - Sound Plurals

Answers by Sound Plural Type

- i
- ijet
- iet
- a
- at
- in
- s
- ien
- ejn
- n
- an

Sound Plural Types

Proportions

0.0  0.1  0.2  0.3  0.4  0.5
Maltese Experiment
Results - Sound Plurals

- *-i* and *-ijiet* are the most common suffixes in our corpus, too

One participant of the experiment said:

„When we [=the Maltese native speakers] do not know the word, we just put an *-i* or *-ijiet* on it. That will leave the word as it is and we avoid mistakes.“
Maltese Experiment
Results - Broken Plurals

Answers by Broken Plural Type

Proportions
Maltese Experiment
Results - Broken Plurals

<table>
<thead>
<tr>
<th>Patterns</th>
<th>Wug Words (sg.-pl.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCVVC</td>
<td>telleb – tlieb</td>
</tr>
<tr>
<td>CCVVCVC</td>
<td>pežna - pżieżeń</td>
</tr>
<tr>
<td>CVCVC</td>
<td>bačča - bačeć</td>
</tr>
</tbody>
</table>

Table 1: Most frequent broken plural patterns in our data

According to Schembri (2012) these patterns are highly productive in Maltese
### Table 2: Proportion of non-canonical plural forms for existing singular nouns

<table>
<thead>
<tr>
<th></th>
<th>Non-canonical frequent</th>
<th>Non-canonical infrequent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sound</td>
<td>Broken</td>
</tr>
<tr>
<td></td>
<td>5(of 400)</td>
<td>1(of 400)</td>
</tr>
<tr>
<td></td>
<td>1,3%</td>
<td>0,3%</td>
</tr>
</tbody>
</table>

- Non-canonical plural forms = forms we do not find in the dictionary
Summary: Results so far

- Changing consonants and vowels influenced the choice of plural forms.
- The plural form of the existing word used as base for nonce words influenced the choice of plural.
- Participants produced broken plurals for nonce words with the most frequent CV structure, sound plurals for nonce words with most common suffixes.
Naive Discriminative Learning
Baayen (2011), Baayen et al. (2011)

- Computational model of morphological processing
- NDL simulates a learning process
- Supervised learning
- Has been used successfully to model language acquisition (Ramscar, Yarlett, Dye, Denny & Thorpe, 2010)
- Central idea: learning = exploring how events are inter-related, they become associated (see also Plag & Balling (2016))
- inter-related events: Cues and Outcomes
Naive Discriminative Learning
Baayen (2011), Baayen et al. (2011)

- Based on Rescorla-Wagner equations that are well established in cognitive psychology (Rescorla & Wagner, 1972)
- Associations between cues and outcomes at a given time, whereas the strength of an association, the association weight, is defined as follows (Evert & Arppe, 2015):
  - No change if a cue is not present in the input
  - Increased if the cue and outcome co-occur
  - Decreased if the cue occurs without the outcome
- Danks (2003) equilibrium equations: define association strength when a stable state is reached = „adult state of the learner“ (Baayen, 2011)
- Implementation as R package `ndl`
Naive Discriminative Learning
Baayen (2011), Baayen et al. (2011)

Figure 6: Association between Cues and Outcomes
We trained the NDL model on our corpus
We formulated our singular nonce words in bigrams and calculated how the NDL learner would classify them
- Cues: singulars in bigrams, #k – ke - el - lb - b#
- Outcomes: plural types, # k = sound, ke = broken...

The associations between cue and outcome are weighted
We used NDL to predict classification of nonce words
<table>
<thead>
<tr>
<th>Cue</th>
<th>Broken Plural</th>
<th>Sound Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>#k</td>
<td>−0.12</td>
<td>0.62</td>
</tr>
<tr>
<td>ke</td>
<td>0.42</td>
<td>−0.42</td>
</tr>
<tr>
<td>el</td>
<td>0.17</td>
<td>−0.17</td>
</tr>
<tr>
<td>lb</td>
<td>0.17</td>
<td>−0.16</td>
</tr>
<tr>
<td>b#</td>
<td>0.42</td>
<td>0.07</td>
</tr>
<tr>
<td>sum</td>
<td>1.06</td>
<td>−0.06</td>
</tr>
</tbody>
</table>

**Table 3:** Example for NDL association weights predicting outcome „broken“ for singular *kelb*
Modeling our Data: Naive Discriminative Learning

Results

- We compared the classification of participants with NDL.
- NDL correctly classified 65.3% of our observations.

<table>
<thead>
<tr>
<th></th>
<th>broken</th>
<th>sound</th>
</tr>
</thead>
<tbody>
<tr>
<td>broken</td>
<td>0.60</td>
<td>0.40</td>
</tr>
<tr>
<td>sound</td>
<td>0.33</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Table 4: Classification of nonce words by NDL
Let’s compare our results with other models that have been used with Arabic broken plural nouns:


Accuracy of the models ranged between 55.31 – 65.97%

Our NDL analysis: 65.3%
Discussion

- There is structure in our data
- Native speakers are able to inflect novel nouns
- Participants produced more broken plural words when we just changed the vowels of existing singulars to create nonce words
- When both, consonants and vowels, were changed, participants produced the highest number of sound plural forms
- Consonants and vowels are important for the generalizations of broken plurals: evidence for tier separation
- Phonotactics of the singular determines the plural form
- Plurals are generalizable!
Grazzi ṭafna!


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.