NDL Modeling of Maltese Plurals and Intuitions of Native Speakers

Jessica Nieder & Ruben van de Vijver
nieder@phil.hhu.de Ruben.Vijver@hhu.de

DFG Research Unit FOR 2373: Project MALT

2nd Workshop on Spoken Morphology: Phonetics and Phonology of Complex Words,
Schloss Mickeln Düsseldorf, August 23-25, 2017
Maltese Plurals

• 2 main strategies to build the plural of a noun:
  ➢ **Sound Plural**  \( sptar \) – \( sptarijiet \) ‘hospital(s)’
  ➢ **Broken Plural**  \( ballun \) – \( blalen \) ‘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: Predictability

• Is it possible to predict pluralisation of novel words?
• Can novel items be classified as broken or sound plurals?
Predicting Maltese Plurals: Previous accounts

• Farrugia & Rosner (2008) used an artificial neural network to compute broken plural forms on the basis of the classification of Schembri (2012)

• Results: the model did not perform well in generalizing to new forms

• Problem: neural networks usually use many hidden layers → what are consequences for human learning?
Predicting Maltese Plurals: Previous accounts

• More recently: Drake & Sharp (2017) present an account that is based on different implementations of the generalized context model (GCM; see Nosofsky, 1990)

• Results: their best performing restricted GCM model had an accuracy of 77.3% (over 5-fold cross-validation)
  - GCM: classification based on similarity of existing items
  - restricted: classification of test forms only on categories that have the same CV patterns the model was trained on
Predicting Maltese Plurals: Our work

• Previous accounts focus on broken plural prediction only
  ➢ How to account for the choice of plural forms?

• We are using the *Naive Discriminative Learner* introduced by Baayen et al. (2011) to predict both, sound and broken plurals
  ➢ 3 steps: Corpus → Production Experiment → NDL modeling
Maltese Plurals: Hypothesis

• The phonotactics of the singular determines the shape of the plural
• More frequent items are more likely to be generalized than infrequent items.
Maltese Plurals: Corpus

- Corpus of 2369 Maltese singular-plural pairs
- Words were taken from Schembri (2012) and an online corpus by Gatt and Čéplö (2013)
- Checked with Ġabra: online lexicon for Maltese (Camilleri, 2013)
- CV structure
- Corpus frequency number
- Number of syllables

Figure: Distribution of plural types in 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-Gleason, 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: Results

glmer with lme4 package (Bates, Maechler, 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 & Base

Significant difference between List\text{consonantsvowels} and List\text{vowels} (p<0.001)

Significant difference between Base\text{brokenplural} and Base\text{soundplural} (p<0.001)
Maltese Experiment: Results – Sound Plurals

Answers by Sound Plural Type

Sound Plural Types:
- -i
- -ijiet
- -iet
- -a
- -at
- -in
- -s
- -ien
- -ejn
- -n
- -an

Proportions
Maltese Experiment: Results – Broken Plurals

Answers by Broken Plural Type

- CCVVC
- (C)CVCVC
- CCVVCVC
- CCVVVC
- VCCCC
- CCVCVC
- CVCCVC
- CVCCVC
- VVCVC
- VCVC

Proportions
Maltese Experiment: Results – Existing Words

<table>
<thead>
<tr>
<th>Frequent</th>
<th>Infrequent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sound</td>
<td>Broken</td>
</tr>
<tr>
<td>5 (of 400)</td>
<td>1 (of 400)</td>
</tr>
<tr>
<td>1,3%</td>
<td>0,3%</td>
</tr>
</tbody>
</table>

Table: Proportion of errors in plural forms for existing singular nouns

- Error Types: no answer, repetition of singular form, 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: Association between Cues and Outcomes
Modeling our Data: Naive Discriminative Learning

- We trained the NDL model on our corpus
- We formulated our singular items in n-grams (unigrams, bigrams, trigrams) and calculated how the NDL learner would classify them

<table>
<thead>
<tr>
<th>Singulars</th>
<th>Cues</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>experiment ‘experiment’</td>
<td>#e_es_sp_pe_er_ri_im_me_en_nt_t#</td>
<td>sound plural</td>
</tr>
<tr>
<td>barma ‘twist’</td>
<td>#b.ba_ar_rm_ma_a#</td>
<td>sound plural</td>
</tr>
<tr>
<td>tokka ‘pen’</td>
<td>#t_to_ok_kk_ka_a#</td>
<td>broken plural</td>
</tr>
<tr>
<td>midwa ‘clinic’</td>
<td>#m_mi_id_dw_wa_a#</td>
<td>broken plural</td>
</tr>
<tr>
<td>qassis ‘priest’</td>
<td>#q_qa_as_ss_si_is_s#</td>
<td>sound plural</td>
</tr>
</tbody>
</table>

Table: Training data set for the NDL model using bigrams as cues
Modeling our Data: Naive Discriminative Learning

- The associations between cue and outcome are weighted
- We used NDL to predict classification of existing singular forms and 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 (kelb)</td>
<td>1.06</td>
<td>-0.06</td>
</tr>
</tbody>
</table>

Table: Example for NDL association weights predicting outcome „broken“ for singular kelb using bigrams as cues
Modeling our Data: Naive Discriminative Learning

• We compared the classification of participants with the prediction of different cue implementations in NDL

• What implementation best models the intuitions of native speakers on plural formation in Maltese?
Results: NDL model 1 – Unigrams as Cues

<table>
<thead>
<tr>
<th></th>
<th>broken</th>
<th>sound</th>
</tr>
</thead>
<tbody>
<tr>
<td>broken</td>
<td>0.08</td>
<td>0.92</td>
</tr>
<tr>
<td>sound</td>
<td>0.05</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Table: Classification of experimental items by NDL with unigrams as cues

- Very good prediction for sound plurals
- Very poor prediction for broken plurals

example: kelb ‘dog’ → k_e_l_b
Results: NDL model 2 – Bigrams as Cues

<table>
<thead>
<tr>
<th></th>
<th>broken</th>
<th>sound</th>
</tr>
</thead>
<tbody>
<tr>
<td>broken</td>
<td>0.59</td>
<td>0.41</td>
</tr>
<tr>
<td>sound</td>
<td>0.33</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Table: Classification of experimental items by NDL with bigrams as cues

• Acceptable prediction for both plural types
Results: NDL model 3 – Trigrams as Cues

Table: Classification of experimental items by NDL with trigrams as cues

<table>
<thead>
<tr>
<th></th>
<th>broken</th>
<th>sound</th>
</tr>
</thead>
<tbody>
<tr>
<td>broken</td>
<td>0.66</td>
<td>0.34</td>
</tr>
<tr>
<td>sound</td>
<td>0.52</td>
<td>0.48</td>
</tr>
</tbody>
</table>

• Good prediction for broken plurals
• Prediction for sound plurals are chance

Example: kelb ‘dog’ → #ke_kel_elb_lb#
Results: Discussion

• Trigrams are the best predictors for broken plurals – unigrams the worst
• Unigrams are the best predictors for sound plurals – trigrams the worst
  
  ➢ Participants used sound plurals more often and corpus contains more sound plurals: when predicting plural forms with just one element of a word (=unigrams), sound plurals will be the default
  
  ➢ Phonotactics (=trigrams= syllables) is especially important for broken plural predictions
Results: Discussion

• glmer model indicates that changing consonants and vowels influenced the choice of plural forms

• Can the NDL model capture this?

• How important are consonants and vowels for the NDL model?

• We changed vowels in cues to V, consonants to C to delete vowel and consonant identity:

  - barma ‘twist’ → #b_bV_Vr_rm_mV_V#
  - barma ‘twist’ → #C_Ca_aC_CC_Ca_a#
### Results: NDL models – Vowels as V

<table>
<thead>
<tr>
<th></th>
<th>broken</th>
<th>sound</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>unigrams</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>broken</td>
<td>0.13</td>
<td>0.87</td>
</tr>
<tr>
<td>sound</td>
<td>0.06</td>
<td>0.94</td>
</tr>
<tr>
<td><strong>bigrams</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>broken</td>
<td>0.39</td>
<td>0.61</td>
</tr>
<tr>
<td>sound</td>
<td>0.25</td>
<td>0.75</td>
</tr>
<tr>
<td><strong>trigrams</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>broken</td>
<td>0.49</td>
<td>0.51</td>
</tr>
<tr>
<td>sound</td>
<td>0.42</td>
<td>0.58</td>
</tr>
</tbody>
</table>

Table: Classification of experimental items with vowels in cues annotated as “V”
Results: NDL models – Consonants as C

<table>
<thead>
<tr>
<th></th>
<th>broken</th>
<th>sound</th>
</tr>
</thead>
<tbody>
<tr>
<td>broken</td>
<td>0.17</td>
<td>0.83</td>
</tr>
<tr>
<td>sound</td>
<td>0.06</td>
<td>0.94</td>
</tr>
</tbody>
</table>

Table: Classification of experimental items with consonants in cues annotated as “C”

example: kelb ‘dog’ → C_e_C_C

example: kelb ‘dog’ → #C_Ce_eC_CC_C#

example: kelb ‘dog’ → #Ce_CeC_eCC_CC#
Results: Discussion

• When all consonants of the experimental items are changed to C we find very poor predictions for broken plurals, regardless of the size of
  gram
  ➢ Consonants are slightly more important for generalization of broken plurals!

• When all vowels of the experimental items are changed to V we find a slightly better performance for broken plurals (especially with bigrams
  and trigrams), nevertheless we cannot replicate the good results of our NDL model 2
  ➢ an abstract representation of consonants and vowels makes the NDL model worse
Modeling our Data: Naive Discriminative Learning

• Let’s compare our results with other models that have been used with Arabic broken plural nouns:

➢ Our best NDL model: 65.3%

Discussion

• Native speakers are able to generalize to novel nouns and use the most common suffixes and CV patterns for this task

• Consonants and vowels are important for the generalizations of Maltese plurals as
  - changing consonants and vowels influenced the choice of plural form of participants and
  - using abstract representations influenced the performance of the NDL models.

• Phonotactics of the singular determines the plural form
Grazzi ħafna!
References


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.