Modeling Maltese broken and sound plurals with Naive Discriminative Learning

Jessica Nieder[a], Fabian Tomaschek[b] & Ruben van de Vijver[a]

nieder@phil.hhu.de, fabian.tomaschek@uni-tuebingen.de, Ruben.Vijver@hhu.de

[a]Department of General Linguistics, Heinrich-Heine-Universität Düsseldorf, Germany;
[b]Department of General Linguistics, Eberhard-Karls-Universität Tübingen, Germany

DFG Research Unit FOR 2373: Project MALT & Project ART

EDLL 2021

Introduction

- How is the content of the mental lexicon structured?
  - **morpheme-based models**: unit of storage = morphemes; we need information about constituent morphemes (e.g. *Item and Arrangement* or *Item and Process*)

    (Hockett, 1954)
  - **word-based models**: unit of storage = words, we need information about the whole word form (e.g. *Word and Paradigm*)

    (Blevins, 2006)
- languages with a rich variety of inflections are a challenge for the prediction of inflectional classes → Maltese!

- 2 main ways to characterize a plural of a noun:
  12 sound plural suffixes: concatenative (traditionally called suffixation)
  annimal - *annimali* 'animal(s)'
  11 broken plural patterns: non-concatenative as changes in the syllabic structure of the plural in comparison to the singular
  ballun - *blalen* 'ball(s)'

Research Question

Is it possible to computationally model complex word forms without morphemes?

This talk

- computational modeling of the Maltese noun plural system without morphemes using the Naive Discriminative Learner (NDL)

Experimental Background

- production experiment with existing singular nouns and phonotactically legal nonce-singulars
- Maltese speakers produced plurals for given singulars
- results: a positive correlation between produced plural forms and distribution of plural forms in the data set
- Maltese speakers use most frequent sound suffixes and broken plural pattern to pluralize novel items: *-i* (41%), *-iet* (23%),*-ijiet* (22%), *CCVVCVC* (33%), *(C)CVCVC* (23%) and *CCVVC* (19%)

(Nieder, van de Vijver, & Mitterer, 2020)

Modeling Maltese Plurals

Data set

- models are based on a data set of 3190 Maltese singular-plural pairs (2406 sound nouns, 784 broken nouns)
- manually transcribed the singular-plural pairs such that every phone is represented as exactly one letter or symbol

NDL Classification

- implementation as R package *ndl*

(Arppe, Hendrix, Milin, et al., 2015; R Core Team, 2019)

- based on discriminative learning
- implements error-driven learning rule

(Rescorla & Wagner, 1972)

- central idea: learning = exploring how events are inter-related, how they become associated

(Plag & Balling, 2016)

- inter-related events: *cues* and *outcomes* (two-layer network)

- validation set approach: randomly divided corpus into a training data set (90%) and a test data set (10%)
- cues = a) singulars only, b) plurals only and c) the whole paradigm (a combination of singular and plural forms) coded as single phones, diphones, triphones
- outcomes = sound vs. broken

Modeling Maltese Plurals

Naive Discriminative Learning - Baseline model

- cue structure:

Test Class	**1phone**	**2phones**	**3phones**
singular	k_e_l_b	#k_ke_el_lb_b#	#ke_kel_elb_lb#
plural	k_l_i:_b	#k_kl_li:_i:b_b#	#kl_kli:_li:b_i:b#
singular-plural	k_e_l_b_k_l_i:_b	#k_ke_el_lb_b#_#k_kl_li:_i:b_b#	#ke_kel_elb_lb#_#kl_kli:_li:b_i:b#

Modeling Maltese Plurals

Naive Discriminative Learning - Baseline model

	good prediction					
CUES	**1phone**		**2phones**		**3phones**	
	broken	sound	broken	sound	broken	sound
singular	3%	99%	24%	86%	39%	67%
plural	61%	97%	89%	98%	64%	87%
singular - plural	47%	94%	97%	68%	76%	56%

Table 1: accuracies for outcomes *sound* vs. *broken* of NDL models with 9 different cue structures: a) singular, b) plural, c) paradigm coded as 1) 1phone, 2) 2phones or 3) 3phones

- best predictions when plurals only or the whole paradigm is used as cues $\rightarrow$ information about the plural is necessary for correct predictions!
- results are in line with a word-based approach of morphology

Modeling Maltese Plurals

Things to think about...

- we simplified Maltese considerably and modeled only two outcomes (sound and broken) BUT Maltese is more complicated
- Nieder, van de Vijver, and Mitterer (2020) and Nieder, van de Vijver, and Mitterer (accepted) found that the frequency of suffixes and patterns plays an important role for morphological processing in Maltese
- How to deal with that?
- solution: include more plural types and their frequencies in the model
- is NDL still able to predict specific patterns/suffixes?

Modeling Maltese Plurals

Data set

- based on experimental background we distinguished 8 plural classes:

1. *-i*
2. *-iet*
3. *-ijiet*
4. *sound* (all other infrequent sound plurals)
5. *broken A, CCVVCVC*
6. *broken B, (C)CVCVC*
7. *broken C, CCVVC*
8. *broken* (all other infrequent broken plurals)

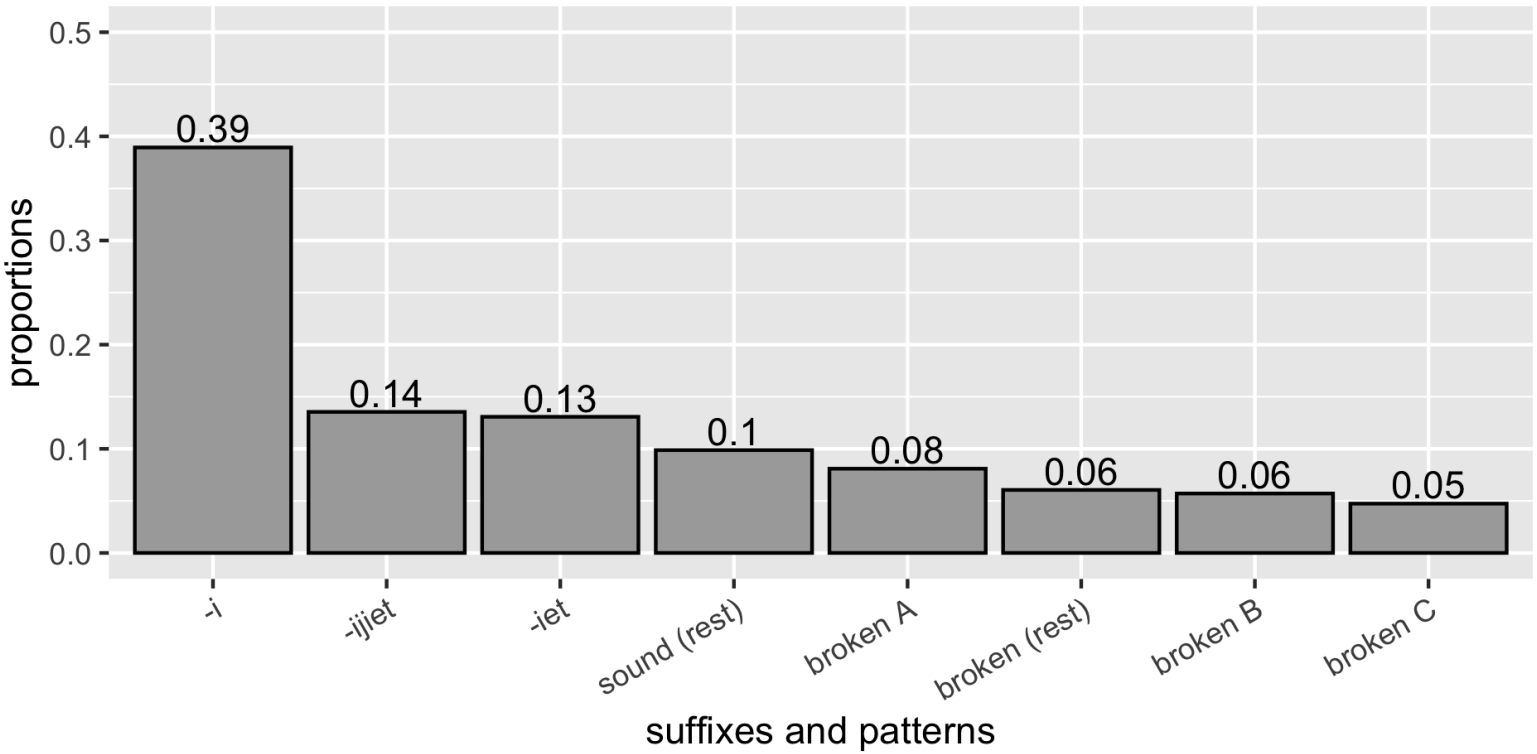

Figure 1: suffixes and patterns in the data set

Modeling Maltese Plurals

Naive Discriminative Learning - pattern/suffix models

- we tested 6 different cue structures:
  1. singulars as 2-phones
  2. singulars as 3-phones
  3. plurals as 2-phones
  4. plurals as 3-phones
  5. singular-plural pairs as 2-phones
  6. singular-plural pairs as 3-phones
- outcomes = 8 plural classes (= chance level at 12.5%)

Modeling Maltese Plurals

NDL - singulars as diphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**8** (38%)	1 (5%)	1 (5%)	0 (0%)	2 (10%)	0 (0%)	7 (33%)	2 (10%)
broken B	0 (0%)	**5** (22%)	1 (4%)	1 (4%)	3 (13%)	2 (9%)	10 (43%)	1 (4%)
broken C	0 (0%)	1 (8%)	**2** (17%)	0 (0%)	4 (33%)	2 (17%)	2 (17%)	1 (8%)
broken (rest)	0 (0%)	0 (0%)	2 (10%)	**4** (19%)	5 (24%)	0 (0%)	6 (29%)	4 (19%)
sound iet	1 (2%)	2 (5%)	3 (8%)	4 (10%)	**16** (40%)	4 (10%)	8 (20%)	2 (5%)
sound ijiet	0 (0%)	0 (0%)	0 (0%)	1 (2%)	0 (0%)	**26** (63%)	12 (29%)	2 (5%)
sound i	3 (2%)	1 (1%)	2 (1%)	0 (0%)	2 (1%)	4 (3%)	**124** (91%)	1 (1%)
sound (rest)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (4%)	4 (17%)	3 (12%)	**16** (67%)

Table 2: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **2-phones of singulars as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 63.01%, accuracy of the worst fold was 56.43%.

Modeling Maltese Plurals

NDL - singulars as diphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**8** (38%)	1 (5%)	1 (5%)	0 (0%)	2 (10%)	0 (0%)	7 (33%)	2 (10%)
broken B	0 (0%)	**5** (22%)	1 (4%)	1 (4%)	3 (13%)	2 (9%)	10 (43%)	1 (4%)
broken C	0 (0%)	1 (8%)	**2** (17%)	0 (0%)	4 (33%)	2 (17%)	2 (17%)	1 (8%)
broken (rest)	0 (0%)	0 (0%)	2 (10%)	**4** (19%)	5 (24%)	0 (0%)	6 (29%)	4 (19%)
sound iet	1 (2%)	2 (5%)	3 (8%)	4 (10%)	**16** (40%)	4 (10%)	8 (20%)	2 (5%)
sound ijiet	0 (0%)	0 (0%)	0 (0%)	1 (2%)	0 (0%)	**26** (63%)	12 (29%)	2 (5%)
sound i	3 (2%)	1 (1%)	2 (1%)	0 (0%)	2 (1%)	4 (3%)	**124** (91%)	1 (1%)
sound (rest)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (4%)	4 (17%)	3 (12%)	**16** (67%)

Table 3: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **2-phones of singulars as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 63.01%, accuracy of the worst fold was 56.43%.

Modeling Maltese Plurals

NDL - singulars as diphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**8** (38%)	1 (5%)	1 (5%)	0 (0%)	2 (10%)	0 (0%)	7 (33%)	2 (10%)
broken B	0 (0%)	**5** (22%)	1 (4%)	1 (4%)	3 (13%)	2 (9%)	10 (43%)	1 (4%)
broken C	0 (0%)	1 (8%)	**2** (17%)	0 (0%)	4 (33%)	2 (17%)	2 (17%)	1 (8%)
broken (rest)	0 (0%)	0 (0%)	2 (10%)	**4** (19%)	5 (24%)	0 (0%)	6 (29%)	4 (19%)
sound iet	1 (2%)	2 (5%)	3 (8%)	4 (10%)	**16** (40%)	4 (10%)	8 (20%)	2 (5%)
sound ijiet	0 (0%)	0 (0%)	0 (0%)	1 (2%)	0 (0%)	**26** (63%)	12 (29%)	2 (5%)
sound i	3 (2%)	1 (1%)	2 (1%)	0 (0%)	2 (1%)	4 (3%)	**124** (91%)	1 (1%)
sound (rest)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (4%)	4 (17%)	3 (12%)	**16** (67%)

Table 4: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **2-phones of singulars as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 63.01%, accuracy of the worst fold was 56.43%.

Modeling Maltese Plurals

NDL - singulars as diphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**8** (38%)	1 (5%)	1 (5%)	0 (0%)	2 (10%)	0 (0%)	7 (33%)	2 (10%)
broken B	0 (0%)	**5** (22%)	1 (4%)	1 (4%)	3 (13%)	2 (9%)	10 (43%)	1 (4%)
broken C	0 (0%)	1 (8%)	**2** (17%)	0 (0%)	4 (33%)	2 (17%)	2 (17%)	1 (8%)
broken (rest)	0 (0%)	0 (0%)	2 (10%)	**4** (19%)	5 (24%)	0 (0%)	6 (29%)	4 (19%)
sound iet	1 (2%)	2 (5%)	3 (8%)	4 (10%)	**16** (40%)	4 (10%)	8 (20%)	2 (5%)
sound ijiet	0 (0%)	0 (0%)	0 (0%)	1 (2%)	0 (0%)	**26** (63%)	12 (29%)	2 (5%)
sound i	3 (2%)	1 (1%)	2 (1%)	0 (0%)	2 (1%)	4 (3%)	**124** (91%)	1 (1%)
sound (rest)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (4%)	4 (17%)	3 (12%)	**16** (67%)

Table 5: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **2-phones of singulars as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 63.01%, accuracy of the worst fold was 56.43%.

Modeling Maltese Plurals

NDL - singulars as diphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**8** (38%)	1 (5%)	1 (5%)	0 (0%)	2 (10%)	0 (0%)	7 (33%)	2 (10%)
broken B	0 (0%)	**5** (22%)	1 (4%)	1 (4%)	3 (13%)	2 (9%)	10 (43%)	1 (4%)
broken C	0 (0%)	1 (8%)	**2** (17%)	0 (0%)	4 (33%)	2 (17%)	2 (17%)	1 (8%)
broken (rest)	0 (0%)	0 (0%)	2 (10%)	**4** (19%)	5 (24%)	0 (0%)	6 (29%)	4 (19%)
sound iet	1 (2%)	2 (5%)	3 (8%)	4 (10%)	**16** (40%)	4 (10%)	8 (20%)	2 (5%)
sound ijiet	0 (0%)	0 (0%)	0 (0%)	1 (2%)	0 (0%)	**26** (63%)	12 (29%)	2 (5%)
sound i	3 (2%)	1 (1%)	2 (1%)	0 (0%)	2 (1%)	4 (3%)	**124** (91%)	1 (1%)
sound (rest)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (4%)	4 (17%)	3 (12%)	**16** (67%)

Table 6: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **2-phones of singulars as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 63.01%, accuracy of the worst fold was 56.43%.

Modeling Maltese Plurals

NDL - singulars as diphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**8** (38%)	1 (5%)	1 (5%)	0 (0%)	2 (10%)	0 (0%)	7 (33%)	2 (10%)
broken B	0 (0%)	**5** (22%)	1 (4%)	1 (4%)	3 (13%)	2 (9%)	10 (43%)	1 (4%)
broken C	0 (0%)	1 (8%)	**2** (17%)	0 (0%)	4 (33%)	2 (17%)	2 (17%)	1 (8%)
broken (rest)	0 (0%)	0 (0%)	2 (10%)	**4** (19%)	5 (24%)	0 (0%)	6 (29%)	4 (19%)
sound iet	1 (2%)	2 (5%)	3 (8%)	4 (10%)	**16** (40%)	4 (10%)	8 (20%)	2 (5%)
sound ijiet	0 (0%)	0 (0%)	0 (0%)	1 (2%)	0 (0%)	**26** (63%)	12 (29%)	2 (5%)
sound i	3 (2%)	1 (1%)	2 (1%)	0 (0%)	2 (1%)	4 (3%)	**124** (91%)	1 (1%)
sound (rest)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (4%)	4 (17%)	3 (12%)	**16** (67%)

Table 7: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **2-phones of singulars as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 63.01%, accuracy of the worst fold was 56.43%.

- providing NDL with information about singulars only results in worse predictions for broken plurals than for sound plurals

Modeling Maltese Plurals

NDL - singulars as triphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**11** (35%)	3 (10%)	0 (0%)	2 (6%)	5 (16%)	0 (0%)	7 (23%)	3 (10%)
broken B	1 (6%)	**5** (28%)	4 (22%)	0 (0%)	4 (22%)	0 (0%)	1 (6%)	3 (17%)
broken C	1 (5%)	2 (11%)	**4** (21%)	1 (5%)	5 (26%)	4 (21%)	2 (11%)	0 (0%)
broken (rest)	0 (0%)	1 (7%)	2 (14%)	**6** (43%)	0 (0%)	2 (14%)	1 (7%)	2 (14%)
sound iet	1 (2%)	1 (2%)	3 (7%)	6 (14%)	**21** (48%)	1 (2%)	11 (25%)	0 (0%)
sound ijiet	0 (0%)	2 (5%)	0 (0%)	1 (3%)	3 (8%)	**19** (51%)	10 (27%)	2 (5%)
sound i	8 (6%)	8 (6%)	3 (2%)	9 (7%)	12 (9%)	11 (9%)	**70** (55%)	7 (5%)
sound (rest)	2 (7%)	1 (4%)	0 (0%)	3 (11%)	0 (0%)	2 (7%)	4 (14%)	**16** (57%)

Table 8: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **3-phones of singulars as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 47.65%, accuracy of the worst fold was 36.05%.

Modeling Maltese Plurals

NDL - singulars as triphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**11** (35%)	3 (10%)	0 (0%)	2 (6%)	5 (16%)	0 (0%)	7 (23%)	3 (10%)
broken B	1 (6%)	**5** (28%)	4 (22%)	0 (0%)	4 (22%)	0 (0%)	1 (6%)	3 (17%)
broken C	1 (5%)	2 (11%)	**4** (21%)	1 (5%)	5 (26%)	4 (21%)	2 (11%)	0 (0%)
broken (rest)	0 (0%)	1 (7%)	2 (14%)	**6** (43%)	0 (0%)	2 (14%)	1 (7%)	2 (14%)
sound iet	1 (2%)	1 (2%)	3 (7%)	6 (14%)	**21** (48%)	1 (2%)	11 (25%)	0 (0%)
sound ijiet	0 (0%)	2 (5%)	0 (0%)	1 (3%)	3 (8%)	**19** (51%)	10 (27%)	2 (5%)
sound i	8 (6%)	8 (6%)	3 (2%)	9 (7%)	12 (9%)	11 (9%)	**70** (55%)	7 (5%)
sound (rest)	2 (7%)	1 (4%)	0 (0%)	3 (11%)	0 (0%)	2 (7%)	4 (14%)	**16** (57%)

Table 9: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **3-phones of singulars as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 47.65%, accuracy of the worst fold was 36.05%.

Modeling Maltese Plurals

NDL - singulars as triphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**11** (35%)	3 (10%)	0 (0%)	2 (6%)	5 (16%)	0 (0%)	7 (23%)	3 (10%)
broken B	1 (6%)	**5** (28%)	4 (22%)	0 (0%)	4 (22%)	0 (0%)	1 (6%)	3 (17%)
broken C	1 (5%)	2 (11%)	**4** (21%)	1 (5%)	5 (26%)	4 (21%)	2 (11%)	0 (0%)
broken (rest)	0 (0%)	1 (7%)	2 (14%)	**6** (43%)	0 (0%)	2 (14%)	1 (7%)	2 (14%)
sound iet	1 (2%)	1 (2%)	3 (7%)	6 (14%)	**21** (48%)	1 (2%)	11 (25%)	0 (0%)
sound ijiet	0 (0%)	2 (5%)	0 (0%)	1 (3%)	3 (8%)	**19** (51%)	10 (27%)	2 (5%)
sound i	8 (6%)	8 (6%)	3 (2%)	9 (7%)	12 (9%)	11 (9%)	**70** (55%)	7 (5%)
sound (rest)	2 (7%)	1 (4%)	0 (0%)	3 (11%)	0 (0%)	2 (7%)	4 (14%)	**16** (57%)

Table 10: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **3-phones of singulars as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 47.65%, accuracy of the worst fold was 36.05%.

Modeling Maltese Plurals

NDL - singulars as triphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**11** (35%)	3 (10%)	0 (0%)	2 (6%)	5 (16%)	0 (0%)	7 (23%)	3 (10%)
broken B	1 (6%)	**5** (28%)	4 (22%)	0 (0%)	4 (22%)	0 (0%)	1 (6%)	3 (17%)
broken C	1 (5%)	2 (11%)	**4** (21%)	1 (5%)	5 (26%)	4 (21%)	2 (11%)	0 (0%)
broken (rest)	0 (0%)	1 (7%)	2 (14%)	**6** (43%)	0 (0%)	2 (14%)	1 (7%)	2 (14%)
sound iet	1 (2%)	1 (2%)	3 (7%)	6 (14%)	**21** (48%)	1 (2%)	11 (25%)	0 (0%)
sound ijiet	0 (0%)	2 (5%)	0 (0%)	1 (3%)	3 (8%)	**19** (51%)	10 (27%)	2 (5%)
sound i	8 (6%)	8 (6%)	3 (2%)	9 (7%)	12 (9%)	11 (9%)	**70** (55%)	7 (5%)
sound (rest)	2 (7%)	1 (4%)	0 (0%)	3 (11%)	0 (0%)	2 (7%)	4 (14%)	**16** (57%)

Table 11: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **3-phones of singulars as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 47.65%, accuracy of the worst fold was 36.05%.

Modeling Maltese Plurals

NDL - singulars as triphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**11** (35%)	3 (10%)	0 (0%)	2 (6%)	5 (16%)	0 (0%)	7 (23%)	3 (10%)
broken B	1 (6%)	**5** (28%)	4 (22%)	0 (0%)	4 (22%)	0 (0%)	1 (6%)	3 (17%)
broken C	1 (5%)	2 (11%)	**4** (21%)	1 (5%)	5 (26%)	4 (21%)	2 (11%)	0 (0%)
broken (rest)	0 (0%)	1 (7%)	2 (14%)	**6** (43%)	0 (0%)	2 (14%)	1 (7%)	2 (14%)
sound iet	1 (2%)	1 (2%)	3 (7%)	6 (14%)	**21** (48%)	1 (2%)	11 (25%)	0 (0%)
sound ijiet	0 (0%)	2 (5%)	0 (0%)	1 (3%)	3 (8%)	**19** (51%)	10 (27%)	2 (5%)
sound i	8 (6%)	8 (6%)	3 (2%)	9 (7%)	12 (9%)	11 (9%)	**70** (55%)	7 (5%)
sound (rest)	2 (7%)	1 (4%)	0 (0%)	3 (11%)	0 (0%)	2 (7%)	4 (14%)	**16** (57%)

Table 12: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **3-phones of singulars as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 47.65%, accuracy of the worst fold was 36.05%.

Modeling Maltese Plurals

NDL - singulars as triphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**11** (35%)	3 (10%)	0 (0%)	2 (6%)	5 (16%)	0 (0%)	7 (23%)	3 (10%)
broken B	1 (6%)	**5** (28%)	4 (22%)	0 (0%)	4 (22%)	0 (0%)	1 (6%)	3 (17%)
broken C	1 (5%)	2 (11%)	**4** (21%)	1 (5%)	5 (26%)	4 (21%)	2 (11%)	0 (0%)
broken (rest)	0 (0%)	1 (7%)	2 (14%)	**6** (43%)	0 (0%)	2 (14%)	1 (7%)	2 (14%)
sound iet	1 (2%)	1 (2%)	3 (7%)	6 (14%)	**21** (48%)	1 (2%)	11 (25%)	0 (0%)
sound ijiet	0 (0%)	2 (5%)	0 (0%)	1 (3%)	3 (8%)	**19** (51%)	10 (27%)	2 (5%)
sound i	8 (6%)	8 (6%)	3 (2%)	9 (7%)	12 (9%)	11 (9%)	**70** (55%)	7 (5%)
sound (rest)	2 (7%)	1 (4%)	0 (0%)	3 (11%)	0 (0%)	2 (7%)	4 (14%)	**16** (57%)

Table 13: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **3-phones of singulars as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 47.65%, accuracy of the worst fold was 36.05%.

- NDL model confuses plurals across categories; less accurate predictions for sound plurals, slightly better prediction for some broken

Modeling Maltese Plurals

NDL - plurals as diphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**23** (74%)	2 (6%)	2 (6%)	1 (3%)	0 (0%)	0 (0%)	0 (0%)	3 (10%)
broken B	2 (13%)	**12** (80%)	1 (7%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
broken C	0 (0%)	2 (15%)	**7** (54%)	0 (0%)	2 (15%)	0 (0%)	0 (0%)	2 (15%)
broken (rest)	0 (0%)	0 (0%)	2 (12%)	**11** (65%)	0 (0%)	0 (0%)	1 (6%)	3 (18%)
sound iet	0 (0%)	0 (0%)	0 (0%)	1 (3%)	**31** (89%)	3 (9%)	0 (0%)	0 (0%)
sound ijiet	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	**45** (100%)	0 (0%)	0 (0%)
sound i	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	**132** (100%)	0 (0%)
sound (rest)	0 (0%)	1 (3%)	0 (0%)	0 (0%)	1 (3%)	0 (0%)	0 (0%)	**29** (94%)

Table 14: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **2-phones of plurals as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 90.91%, accuracy of the worst fold was 84.64%.

Modeling Maltese Plurals

NDL - plurals as diphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**23** (74%)	2 (6%)	2 (6%)	1 (3%)	0 (0%)	0 (0%)	0 (0%)	3 (10%)
broken B	2 (13%)	**12** (80%)	1 (7%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
broken C	0 (0%)	2 (15%)	**7** (54%)	0 (0%)	2 (15%)	0 (0%)	0 (0%)	2 (15%)
broken (rest)	0 (0%)	0 (0%)	2 (12%)	**11** (65%)	0 (0%)	0 (0%)	1 (6%)	3 (18%)
sound iet	0 (0%)	0 (0%)	0 (0%)	1 (3%)	**31** (89%)	3 (9%)	0 (0%)	0 (0%)
sound ijiet	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	**45** (100%)	0 (0%)	0 (0%)
sound i	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	**132** (100%)	0 (0%)
sound (rest)	0 (0%)	1 (3%)	0 (0%)	0 (0%)	1 (3%)	0 (0%)	0 (0%)	**29** (94%)

Table 15: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **2-phones of plurals as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 90.91%, accuracy of the worst fold was 84.64%.

Modeling Maltese Plurals

NDL - plurals as diphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**23** (74%)	2 (6%)	2 (6%)	1 (3%)	0 (0%)	0 (0%)	0 (0%)	3 (10%)
broken B	2 (13%)	**12** (80%)	1 (7%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
broken C	0 (0%)	2 (15%)	**7** (54%)	0 (0%)	2 (15%)	0 (0%)	0 (0%)	2 (15%)
broken (rest)	0 (0%)	0 (0%)	2 (12%)	**11** (65%)	0 (0%)	0 (0%)	1 (6%)	3 (18%)
sound iet	0 (0%)	0 (0%)	0 (0%)	1 (3%)	**31** (89%)	3 (9%)	0 (0%)	0 (0%)
sound ijiet	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	**45** (100%)	0 (0%)	0 (0%)
sound i	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	**132** (100%)	0 (0%)
sound (rest)	0 (0%)	1 (3%)	0 (0%)	0 (0%)	1 (3%)	0 (0%)	0 (0%)	**29** (94%)

Table 16: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **2-phones of plurals as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 90.91%, accuracy of the worst fold was 84.64%.

Modeling Maltese Plurals

NDL - plurals as diphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**23** (74%)	2 (6%)	2 (6%)	1 (3%)	0 (0%)	0 (0%)	0 (0%)	3 (10%)
broken B	2 (13%)	**12** (80%)	1 (7%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
broken C	0 (0%)	2 (15%)	**7** (54%)	0 (0%)	2 (15%)	0 (0%)	0 (0%)	2 (15%)
broken (rest)	0 (0%)	0 (0%)	2 (12%)	**11** (65%)	0 (0%)	0 (0%)	1 (6%)	3 (18%)
sound iet	0 (0%)	0 (0%)	0 (0%)	1 (3%)	**31** (89%)	3 (9%)	0 (0%)	0 (0%)
sound ijiet	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	**45** (100%)	0 (0%)	0 (0%)
sound i	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	**132** (100%)	0 (0%)
sound (rest)	0 (0%)	1 (3%)	0 (0%)	0 (0%)	1 (3%)	0 (0%)	0 (0%)	**29** (94%)

Table 17: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **2-phones of plurals as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 90.91%, accuracy of the worst fold was 84.64%.

Modeling Maltese Plurals

NDL - plurals as diphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**23** (74%)	2 (6%)	2 (6%)	1 (3%)	0 (0%)	0 (0%)	0 (0%)	3 (10%)
broken B	2 (13%)	**12** (80%)	1 (7%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
broken C	0 (0%)	2 (15%)	**7** (54%)	0 (0%)	2 (15%)	0 (0%)	0 (0%)	2 (15%)
broken (rest)	0 (0%)	0 (0%)	2 (12%)	**11** (65%)	0 (0%)	0 (0%)	1 (6%)	3 (18%)
sound iet	0 (0%)	0 (0%)	0 (0%)	1 (3%)	**31** (89%)	3 (9%)	0 (0%)	0 (0%)
sound ijiet	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	**45** (100%)	0 (0%)	0 (0%)
sound i	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	**132** (100%)	0 (0%)
sound (rest)	0 (0%)	1 (3%)	0 (0%)	0 (0%)	1 (3%)	0 (0%)	0 (0%)	**29** (94%)

Table 18: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **2-phones of plurals as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 90.91%, accuracy of the worst fold was 84.64%.

Modeling Maltese Plurals

NDL - plurals as diphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**23** (74%)	2 (6%)	2 (6%)	1 (3%)	0 (0%)	0 (0%)	0 (0%)	3 (10%)
broken B	2 (13%)	**12** (80%)	1 (7%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
broken C	0 (0%)	2 (15%)	**7** (54%)	0 (0%)	2 (15%)	0 (0%)	0 (0%)	2 (15%)
broken (rest)	0 (0%)	0 (0%)	2 (12%)	**11** (65%)	0 (0%)	0 (0%)	1 (6%)	3 (18%)
sound iet	0 (0%)	0 (0%)	0 (0%)	1 (3%)	**31** (89%)	3 (9%)	0 (0%)	0 (0%)
sound ijiet	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	**45** (100%)	0 (0%)	0 (0%)
sound i	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	**132** (100%)	0 (0%)
sound (rest)	0 (0%)	1 (3%)	0 (0%)	0 (0%)	1 (3%)	0 (0%)	0 (0%)	**29** (94%)

Table 19: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **2-phones of plurals as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 90.91%, accuracy of the worst fold was 84.64%.

- excellent predictions for sound, very good predictions for broken

Modeling Maltese Plurals

NDL - plurals as triphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**24** (77%)	1 (3%)	5 (16%)	0 (0%)	0 (0%)	0 (0%)	1 (3%)	0 (0%)
broken B	2 (13%)	**11** (73%)	2 (13%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
broken C	2 (15%)	0 (0%)	**8** (62%)	3 (23%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
broken (rest)	0 (0%)	0 (0%)	1 (6%)	**15** (88%)	0 (0%)	0 (0%)	0 (0%)	1 (6%)
sound iet	0 (0%)	0 (0%)	1 (3%)	0 (0%)	**31** (89%)	3 (9%)	0 (0%)	0 (0%)
sound ijiet	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (4%)	**43** (96%)	0 (0%)	0 (0%)
sound i	1 (1%)	0 (0%)	0 (0%)	6 (5%)	0 (0%)	1 (1%)	**122** (92%)	2 (2%)
sound (rest)	0 (0%)	1 (3%)	0 (0%)	1 (3%)	1 (3%)	0 (0%)	1 (3%)	**27** (87%)

Table 20: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **3-phones of plurals as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 88.09%, accuracy of the worst fold was 74.92%.

Modeling Maltese Plurals

NDL - plurals as triphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**24** (77%)	1 (3%)	5 (16%)	0 (0%)	0 (0%)	0 (0%)	1 (3%)	0 (0%)
broken B	2 (13%)	**11** (73%)	2 (13%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
broken C	2 (15%)	0 (0%)	**8** (62%)	3 (23%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
broken (rest)	0 (0%)	0 (0%)	1 (6%)	**15** (88%)	0 (0%)	0 (0%)	0 (0%)	1 (6%)
sound iet	0 (0%)	0 (0%)	1 (3%)	0 (0%)	**31** (89%)	3 (9%)	0 (0%)	0 (0%)
sound ijiet	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (4%)	**43** (96%)	0 (0%)	0 (0%)
sound i	1 (1%)	0 (0%)	0 (0%)	6 (5%)	0 (0%)	1 (1%)	**122** (92%)	2 (2%)
sound (rest)	0 (0%)	1 (3%)	0 (0%)	1 (3%)	1 (3%)	0 (0%)	1 (3%)	**27** (87%)

Table 21: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **3-phones of plurals as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 88.09%, accuracy of the worst fold was 74.92%.

Modeling Maltese Plurals

NDL - plurals as triphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**24** (77%)	1 (3%)	5 (16%)	0 (0%)	0 (0%)	0 (0%)	1 (3%)	0 (0%)
broken B	2 (13%)	**11** (73%)	2 (13%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
broken C	2 (15%)	0 (0%)	**8** (62%)	3 (23%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
broken (rest)	0 (0%)	0 (0%)	1 (6%)	**15** (88%)	0 (0%)	0 (0%)	0 (0%)	1 (6%)
sound iet	0 (0%)	0 (0%)	1 (3%)	0 (0%)	**31** (89%)	3 (9%)	0 (0%)	0 (0%)
sound ijiet	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (4%)	**43** (96%)	0 (0%)	0 (0%)
sound i	1 (1%)	0 (0%)	0 (0%)	6 (5%)	0 (0%)	1 (1%)	**122** (92%)	2 (2%)
sound (rest)	0 (0%)	1 (3%)	0 (0%)	1 (3%)	1 (3%)	0 (0%)	1 (3%)	**27** (87%)

Table 22: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **3-phones of plurals as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 88.09%, accuracy of the worst fold was 74.92%.

Modeling Maltese Plurals

NDL - plurals as triphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**24** (77%)	1 (3%)	5 (16%)	0 (0%)	0 (0%)	0 (0%)	1 (3%)	0 (0%)
broken B	2 (13%)	**11** (73%)	2 (13%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
broken C	2 (15%)	0 (0%)	**8** (62%)	3 (23%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
broken (rest)	0 (0%)	0 (0%)	1 (6%)	**15** (88%)	0 (0%)	0 (0%)	0 (0%)	1 (6%)
sound iet	0 (0%)	0 (0%)	1 (3%)	0 (0%)	**31** (89%)	3 (9%)	0 (0%)	0 (0%)
sound ijiet	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (4%)	**43** (96%)	0 (0%)	0 (0%)
sound i	1 (1%)	0 (0%)	0 (0%)	6 (5%)	0 (0%)	1 (1%)	**122** (92%)	2 (2%)
sound (rest)	0 (0%)	1 (3%)	0 (0%)	1 (3%)	1 (3%)	0 (0%)	1 (3%)	**27** (87%)

Table 23: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **3-phones of plurals as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 88.09%, accuracy of the worst fold was 74.92%.

Modeling Maltese Plurals

NDL - plurals as triphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**24** (77%)	1 (3%)	5 (16%)	0 (0%)	0 (0%)	0 (0%)	1 (3%)	0 (0%)
broken B	2 (13%)	**11** (73%)	2 (13%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
broken C	2 (15%)	0 (0%)	**8** (62%)	3 (23%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
broken (rest)	0 (0%)	0 (0%)	1 (6%)	**15** (88%)	0 (0%)	0 (0%)	0 (0%)	1 (6%)
sound iet	0 (0%)	0 (0%)	1 (3%)	0 (0%)	**31** (89%)	3 (9%)	0 (0%)	0 (0%)
sound ijiet	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (4%)	**43** (96%)	0 (0%)	0 (0%)
sound i	1 (1%)	0 (0%)	0 (0%)	6 (5%)	0 (0%)	1 (1%)	**122** (92%)	2 (2%)
sound (rest)	0 (0%)	1 (3%)	0 (0%)	1 (3%)	1 (3%)	0 (0%)	1 (3%)	**27** (87%)

Table 24: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **3-phones of plurals as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 88.09%, accuracy of the worst fold was 74.92%.

Modeling Maltese Plurals

NDL - plurals as triphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**24** (77%)	1 (3%)	5 (16%)	0 (0%)	0 (0%)	0 (0%)	1 (3%)	0 (0%)
broken B	2 (13%)	**11** (73%)	2 (13%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
broken C	2 (15%)	0 (0%)	**8** (62%)	3 (23%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
broken (rest)	0 (0%)	0 (0%)	1 (6%)	**15** (88%)	0 (0%)	0 (0%)	0 (0%)	1 (6%)
sound iet	0 (0%)	0 (0%)	1 (3%)	0 (0%)	**31** (89%)	3 (9%)	0 (0%)	0 (0%)
sound ijiet	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (4%)	**43** (96%)	0 (0%)	0 (0%)
sound i	1 (1%)	0 (0%)	0 (0%)	6 (5%)	0 (0%)	1 (1%)	**122** (92%)	2 (2%)
sound (rest)	0 (0%)	1 (3%)	0 (0%)	1 (3%)	1 (3%)	0 (0%)	1 (3%)	**27** (87%)

Table 25: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **3-phones of plurals as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 88.09%, accuracy of the worst fold was 74.92%.

- excellent but worse classification compared to 2-phone cues for sound, better predictions for some broken

Modeling Maltese Plurals

NDL - singulars-plurals as diphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**13** (62%)	2 (10%)	5 (24%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (5%)
broken B	0 (0%)	**17** (74%)	3 (13%)	0 (0%)	0 (0%)	0 (0%)	1 (4%)	2 (9%)
broken C	2 (17%)	1 (8%)	**8** (67%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (8%)
broken (rest)	1 (5%)	2 (10%)	4 (19%)	**6** (29%)	0 (0%)	1 (5%)	3 (14%)	4 (19%)
sound iet	1 (2%)	0 (0%)	2 (5%)	1 (2%)	**29** (72%)	7 (18%)	0 (0%)	0 (0%)
sound ijiet	0 (0%)	0 (0%)	1 (2%)	0 (0%)	0 (0%)	**38** (93%)	2 (5%)	0 (0%)
sound i	1 (1%)	0 (0%)	1 (1%)	0 (0%)	0 (0%)	1 (1%)	**133** (97%)	1 (1%)
sound (rest)	0 (0%)	1 (4%)	2 (8%)	4 (17%)	0 (0%)	0 (0%)	1 (4%)	**16** (67%)

Table 26: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **2-phones of the singular-plural paradigm as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 81.5%, accuracy of the worst fold was 74.29%.

Modeling Maltese Plurals

NDL - singulars-plurals as diphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**13** (62%)	2 (10%)	5 (24%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (5%)
broken B	0 (0%)	**17** (74%)	3 (13%)	0 (0%)	0 (0%)	0 (0%)	1 (4%)	2 (9%)
broken C	2 (17%)	1 (8%)	**8** (67%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (8%)
broken (rest)	1 (5%)	2 (10%)	4 (19%)	**6** (29%)	0 (0%)	1 (5%)	3 (14%)	4 (19%)
sound iet	1 (3%)	1 (3%)	4 (10%)	6 (15%)	**26** (67%)	1 (3%)	0 (0%)	0 (0%)
sound ijiet	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1(2%)	**44** (96%)	1 (2%)	0 (0%)
sound i	0 (0%)	1 (1%)	2 (1%)	4 (3%)	0 (0%)	1 (1%)	**127** (94%)	0 (0%)
sound (rest)	2 (6%)	0 (0%)	5(15%)	3 (9%)	1 (3%)	0 (0%)	6(18%)	**17** (50%)

Table 27: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **2-phones of the singular-plural paradigm as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 81.5%, accuracy of the worst fold was 74.29%.

Modeling Maltese Plurals

NDL - singulars-plurals as diphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**13** (62%)	2 (10%)	5 (24%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (5%)
broken B	0 (0%)	**17** (74%)	3 (13%)	0 (0%)	0 (0%)	0 (0%)	1 (4%)	2 (9%)
broken C	2 (17%)	1 (8%)	**8** (67%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (8%)
broken (rest)	1 (5%)	2 (10%)	4 (19%)	**6** (29%)	0 (0%)	1 (5%)	3 (14%)	4 (19%)
sound iet	1 (3%)	1 (3%)	4 (10%)	6 (15%)	**26** (67%)	1 (3%)	0 (0%)	0 (0%)
sound ijiet	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1(2%)	**44** (96%)	1 (2%)	0 (0%)
sound i	0 (0%)	1 (1%)	2 (1%)	4 (3%)	0 (0%)	1 (1%)	**127** (94%)	0 (0%)
sound (rest)	2 (6%)	0 (0%)	5(15%)	3 (9%)	1 (3%)	0 (0%)	6(18%)	**17** (50%)

Table 28: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **2-phones of the singular-plural paradigm as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 81.5%, accuracy of the worst fold was 74.29%.

Modeling Maltese Plurals

NDL - singulars-plurals as diphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**13** (62%)	2 (10%)	5 (24%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (5%)
broken B	0 (0%)	**17** (74%)	3 (13%)	0 (0%)	0 (0%)	0 (0%)	1 (4%)	2 (9%)
broken C	2 (17%)	1 (8%)	**8** (67%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (8%)
broken (rest)	1 (5%)	2 (10%)	4 (19%)	**6** (29%)	0 (0%)	1 (5%)	3 (14%)	4 (19%)
sound iet	1 (2%)	0 (0%)	2 (5%)	1 (2%)	**29** (72%)	7 (18%)	0 (0%)	0 (0%)
sound ijiet	0 (0%)	0 (0%)	1 (2%)	0 (0%)	0 (0%)	**38** (93%)	2 (5%)	0 (0%)
sound i	1 (1%)	0 (0%)	1 (1%)	0 (0%)	0 (0%)	1 (1%)	**133** (97%)	1 (1%)
sound (rest)	0 (0%)	1 (4%)	2 (8%)	4 (17%)	0 (0%)	0 (0%)	1 (4%)	**16** (67%)

Table 29: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **2-phones of the singular-plural paradigm as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 81.5%, accuracy of the worst fold was 74.29%.

Modeling Maltese Plurals

NDL - singulars-plurals as diphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**13** (62%)	2 (10%)	5 (24%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (5%)
broken B	0 (0%)	**17** (74%)	3 (13%)	0 (0%)	0 (0%)	0 (0%)	1 (4%)	2 (9%)
broken C	2 (17%)	1 (8%)	**8** (67%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (8%)
broken (rest)	1 (5%)	2 (10%)	4 (19%)	**6** (29%)	0 (0%)	1 (5%)	3 (14%)	4 (19%)
sound iet	1 (2%)	0 (0%)	2 (5%)	1 (2%)	**29** (72%)	7 (18%)	0 (0%)	0 (0%)
sound ijiet	0 (0%)	0 (0%)	1 (2%)	0 (0%)	0 (0%)	**38** (93%)	2 (5%)	0 (0%)
sound i	1 (1%)	0 (0%)	1 (1%)	0 (0%)	0 (0%)	1 (1%)	**133** (97%)	1 (1%)
sound (rest)	0 (0%)	1 (4%)	2 (8%)	4 (17%)	0 (0%)	0 (0%)	1 (4%)	**16** (67%)

Table 30: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **2-phones of the singular-plural paradigm as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 81.5%, accuracy of the worst fold was 74.29%.

Modeling Maltese Plurals

NDL - singulars-plurals as diphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**13** (62%)	2 (10%)	5 (24%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (5%)
broken B	0 (0%)	**17** (74%)	3 (13%)	0 (0%)	0 (0%)	0 (0%)	1 (4%)	2 (9%)
broken C	2 (17%)	1 (8%)	**8** (67%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (8%)
broken (rest)	1 (5%)	2 (10%)	4 (19%)	**6** (29%)	0 (0%)	1 (5%)	3 (14%)	4 (19%)
sound iet	1 (2%)	0 (0%)	2 (5%)	1 (2%)	**29** (72%)	7 (18%)	0 (0%)	0 (0%)
sound ijiet	0 (0%)	0 (0%)	1 (2%)	0 (0%)	0 (0%)	**38** (93%)	2 (5%)	0 (0%)
sound i	1 (1%)	0 (0%)	1 (1%)	0 (0%)	0 (0%)	1 (1%)	**133** (97%)	1 (1%)
sound (rest)	0 (0%)	1 (4%)	2 (8%)	4 (17%)	0 (0%)	0 (0%)	1 (4%)	**16** (67%)

Table 31: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **2-phones of the singular-plural paradigm as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 81.5%, accuracy of the worst fold was 74.29%.

- good predictions, confusion mostly within categories

Modeling Maltese Plurals

NDL - singulars-plurals as triphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**19** (61%)	2 (6%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	8 (26%)	2 (6%)
broken B	0 (0%)	**7** (47%)	3 (20%)	1 (7%)	0 (0%)	0 (0%)	3 (20%)	1 (7%)
broken C	0 (0%)	1 (8%)	**3** (23%)	3 (23%)	2 (15%)	0 (0%)	3 (23%)	1 (8%)
broken (rest)	0 (0%)	1 (6%)	1 (6%)	**10** (59%)	0 (0%)	0 (0%)	2 (12%)	3 (18%)
sound iet	0 (0%)	3 (9%)	7 (20%)	2 (6%)	**15** (43%)	1 (3%)	5 (14%)	2 (6%)
sound ijiet	0 (0%)	2 (4%)	5 (11%)	1 (2%)	0 (0%)	**17** (38%)	19 (42%)	1 (2%)
sound i	0 (0%)	9 (7%)	26 (20%)	3 (2%)	0 (0%)	0 (0%)	**79** (60%)	15 (11%)
sound (rest)	1 (3%)	2 (6%)	4 (13%)	1 (3%)	1 (3%)	0 (0%)	3 (10%)	**19** (61%)

Table 32: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **3-phones of the singular-plural paradigm as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 52.98%, accuracy of the worst fold was 30.41%.

Modeling Maltese Plurals

NDL - singulars-plurals as triphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**19** (61%)	2 (6%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	8 (26%)	2 (6%)
broken B	0 (0%)	**7** (47%)	3 (20%)	1 (7%)	0 (0%)	0 (0%)	3 (20%)	1 (7%)
broken C	0 (0%)	1 (8%)	**3** (23%)	3 (23%)	2 (15%)	0 (0%)	3 (23%)	1 (8%)
broken (rest)	0 (0%)	1 (6%)	1 (6%)	**10** (59%)	0 (0%)	0 (0%)	2 (12%)	3 (18%)
sound iet	0 (0%)	3 (9%)	7 (20%)	2 (6%)	**15** (43%)	1 (3%)	5 (14%)	2 (6%)
sound ijiet	0 (0%)	2 (4%)	5 (11%)	1 (2%)	0 (0%)	**17** (38%)	19 (42%)	1 (2%)
sound i	0 (0%)	9 (7%)	26 (20%)	3 (2%)	0 (0%)	0 (0%)	**79** (60%)	15 (11%)
sound (rest)	1 (3%)	2 (6%)	4 (13%)	1 (3%)	1 (3%)	0 (0%)	3 (10%)	**19** (61%)

Table 33: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **3-phones of the singular-plural paradigm as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 52.98%, accuracy of the worst fold was 30.41%.

Modeling Maltese Plurals

NDL - singulars-plurals as triphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**19** (61%)	2 (6%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	8 (26%)	2 (6%)
broken B	0 (0%)	**7** (47%)	3 (20%)	1 (7%)	0 (0%)	0 (0%)	3 (20%)	1 (7%)
broken C	0 (0%)	1 (8%)	**3** (23%)	3 (23%)	2 (15%)	0 (0%)	3 (23%)	1 (8%)
broken (rest)	0 (0%)	1 (6%)	1 (6%)	**10** (59%)	0 (0%)	0 (0%)	2 (12%)	3 (18%)
sound iet	0 (0%)	3 (9%)	7 (20%)	2 (6%)	**15** (43%)	1 (3%)	5 (14%)	2 (6%)
sound ijiet	0 (0%)	2 (4%)	5 (11%)	1 (2%)	0 (0%)	**17** (38%)	19 (42%)	1 (2%)
sound i	0 (0%)	9 (7%)	26 (20%)	3 (2%)	0 (0%)	0 (0%)	**79** (60%)	15 (11%)
sound (rest)	1 (3%)	2 (6%)	4 (13%)	1 (3%)	1 (3%)	0 (0%)	3 (10%)	**19** (61%)

Table 34: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **3-phones of the singular-plural paradigm as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 52.98%, accuracy of the worst fold was 30.41%.

Modeling Maltese Plurals

NDL - singulars-plurals as triphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**19** (61%)	2 (6%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	8 (26%)	2 (6%)
broken B	0 (0%)	**7** (47%)	3 (20%)	1 (7%)	0 (0%)	0 (0%)	3 (20%)	1 (7%)
broken C	0 (0%)	1 (8%)	**3** (23%)	3 (23%)	2 (15%)	0 (0%)	3 (23%)	1 (8%)
broken (rest)	0 (0%)	1 (6%)	1 (6%)	**10** (59%)	0 (0%)	0 (0%)	2 (12%)	3 (18%)
sound iet	0 (0%)	3 (9%)	7 (20%)	2 (6%)	**15** (43%)	1 (3%)	5 (14%)	2 (6%)
sound ijiet	0 (0%)	2 (4%)	5 (11%)	1 (2%)	0 (0%)	**17** (38%)	19 (42%)	1 (2%)
sound i	0 (0%)	9 (7%)	26 (20%)	3 (2%)	0 (0%)	0 (0%)	**79** (60%)	15 (11%)
sound (rest)	1 (3%)	2 (6%)	4 (13%)	1 (3%)	1 (3%)	0 (0%)	3 (10%)	**19** (61%)

Table 35: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **3-phones of the singular-plural paradigm as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 52.98%, accuracy of the worst fold was 30.41%.

Modeling Maltese Plurals

NDL - singulars-plurals as triphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**19** (61%)	2 (6%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	8 (26%)	2 (6%)
broken B	0 (0%)	**7** (47%)	3 (20%)	1 (7%)	0 (0%)	0 (0%)	3 (20%)	1 (7%)
broken C	0 (0%)	1 (8%)	**3** (23%)	3 (23%)	2 (15%)	0 (0%)	3 (23%)	1 (8%)
broken (rest)	0 (0%)	1 (6%)	1 (6%)	**10** (59%)	0 (0%)	0 (0%)	2 (12%)	3 (18%)
sound iet	0 (0%)	3 (9%)	7 (20%)	2 (6%)	**15** (43%)	1 (3%)	5 (14%)	2 (6%)
sound ijiet	0 (0%)	2 (4%)	5 (11%)	1 (2%)	0 (0%)	**17** (38%)	19 (42%)	1 (2%)
sound i	0 (0%)	9 (7%)	26 (20%)	3 (2%)	0 (0%)	0 (0%)	**79** (60%)	15 (11%)
sound (rest)	1 (3%)	2 (6%)	4 (13%)	1 (3%)	1 (3%)	0 (0%)	3 (10%)	**19** (61%)

Table 36: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **3-phones of the singular-plural paradigm as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 52.98%, accuracy of the worst fold was 30.41%.

Modeling Maltese Plurals

NDL - singulars-plurals as triphone cues

	broken A	broken B	broken C	broken	sound iet	sound ijiet	sound i	sound
broken A	**19** (61%)	2 (6%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	8 (26%)	2 (6%)
broken B	0 (0%)	**7** (47%)	3 (20%)	1 (7%)	0 (0%)	0 (0%)	3 (20%)	1 (7%)
broken C	0 (0%)	1 (8%)	**3** (23%)	3 (23%)	2 (15%)	0 (0%)	3 (23%)	1 (8%)
broken (rest)	0 (0%)	1 (6%)	1 (6%)	**10** (59%)	0 (0%)	0 (0%)	2 (12%)	3 (18%)
sound iet	0 (0%)	3 (9%)	7 (20%)	2 (6%)	**15** (43%)	1 (3%)	5 (14%)	2 (6%)
sound ijiet	0 (0%)	2 (4%)	5 (11%)	1 (2%)	0 (0%)	**17** (38%)	19 (42%)	1 (2%)
sound i	0 (0%)	9 (7%)	26 (20%)	3 (2%)	0 (0%)	0 (0%)	**79** (60%)	15 (11%)
sound (rest)	1 (3%)	2 (6%)	4 (13%)	1 (3%)	1 (3%)	0 (0%)	3 (10%)	**19** (61%)

Table 37: Confusion matrix of the best fold in 10 fold cross-validation of having the NDL model predicting plural class with **3-phones of the singular-plural paradigm as cues**. Rows represent the input category, columns represent their classification. Accuracy of the best fold was 52.98%, accuracy of the worst fold was 30.41%.

- overall less accurate predictions

Modeling Maltese Plurals

NDL: Conclusion

- information about the plurals is necessary for correct predictions
- 2-phone cues are more informative than 3-phone cues for sound, for broken 3-phone cues are more informative $\rightarrow$ two morphological systems based on different information
- best 3 models: models with the plural coded as 2-phone and 3-phone cues and model with singular-plural paradigm coded as 2-phone cues (accuracy of the best folds: 90.91%, 88.09%, and 81.5%)

Discussion & Conclusion

- we classified Maltese plurals with several NDL models $\rightarrow$ good predictions for broken vs. sound and for 8 plural classes
- we don't need morphemes for classification of Maltese plurals (but we need information about the plural forms)
- our results are in line with a word-based model of morphological processing

Grazzi ħafna!

References I

Arppe, A., Hendrix, P., Milin, P., Baayen, R. H., Sering, T., & Shaoul, C. (2015). *Ndl: Naive Discriminative Learning*. https://CRAN.R-project.org/package=ndl.

Blevins, J. P. (2006). Word-based morphology. *Journal of Linguistics*, *42*, 531–573.

Hockett, C. F. (1954). Two models of grammatical description. *Word*, *10*(2-3), 210–234.

Nieder, J., van de Vijver, R., & Mitterer, H. (2020). Knowledge of Maltese singular-plural mappings. Analogy explains it best. *Morphology*. https://doi.org/https://doi.org/10.1007/s11525-020-09353-7

Nieder, J., van de Vijver, R., & Mitterer, H. (accepted). Priming Maltese Plurals: Representation of sound and broken plurals in the mental lexicon. To appear in: *Mental Lexicon*, *16*(1).

References II

Plag, I., & Balling, L. W. (2016). Derivational morphology: An integrative perspective on some fundamental issues. In V. Pirelli, I. Plag, & W. U. Dressler (Eds.), *Word knowledge and word usage: A cross-disciplinary guide to the mental lexicon*. De Gruyter.

R Core Team. (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.

Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. Black & W. Prokasy (Eds.), *Classical conditioning ii: Current research and theory* (pp. 64–99). Appleton-Century-Crofts.