

# Predicting Maltese Plural Patterns with Naive Discriminative Learning

Jessica Nieder & Ruben van de Vijver & Fabian Tomaschek

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

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

Words in the World International Conference, 16-18 October 2020

- what information do native speakers use to inflect novel word forms?
  - **compositional view of complex words**: we need information about constituent morphemes  
(Zwitserslood, 2018)
  - **word and paradigm morphology**: we need information about the whole word form and its paradigm  
(Blevins, 2016)
- 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 patterns:** concatenative (traditionally called suffixation)  
*animal - animali* '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)'

# Naive Discriminative Learning

- implementation as R package *ndl*  
(Arppe, Hendrix, Milin, et al., 2015; R Core Team, 2019)
- based on discriminative learning  
(Ramscar, Dye, & McCauley, 2013; Ramscar, Yarlett, Dye, et al., 2010)
- 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)

# Naive Discriminative Learning

- associations between cues and outcomes at a given time, whereas the strength of an association, the association weight, is defined as follows:
  - 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

(Evert & Arppe, 2015)
- Danks, 2003 equilibrium equations: define association strength when a stable state is reached = "adult state of the learner"

(Baayen, 2011)

# Naive Discriminative Learning

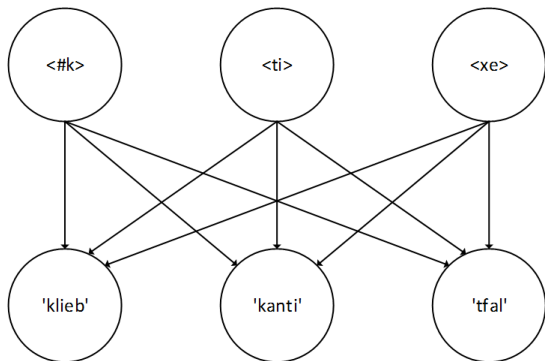


Figure 1: association between cues and outcomes

- is it possible to model classification and production of the Maltese noun plural system without using morphemes?

# Experimental Background

- production experiment on Maltese plurals 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

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



# Modeling Maltese Plurals

- NDL 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 phoneme is represented as exactly one letter or symbol
- validation set approach: randomly divided corpus into a training data set (90%) and a test data set (10%)
- as cues we used a) singulars only, b) plurals only and c) the whole paradigm (a combination of singular and plural forms)
- cues were coded as single phones, diphones or triphones
- outcomes are the categories sound and broken

# Modeling Maltese Plurals

| CUES                     | 1phone |       | 2phones |       | 3phones |       |
|--------------------------|--------|-------|---------|-------|---------|-------|
|                          | broken | sound | broken  | sound | broken  | sound |
| <b>singular</b>          | 3%     | 99%   | 24%     | 86%   | 39%     | 67%   |
| <b>plural</b>            | 61%    | 97%   | 89%     | 98%   | 64%     | 87%   |
| <b>singular - plural</b> | 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

# Modeling Maltese Plurals

| 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 2:** 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

# Modeling Maltese Plurals

| CUES                     | 1phone |       | 2phones |       | 3phones |       |
|--------------------------|--------|-------|---------|-------|---------|-------|
|                          | broken | sound | broken  | sound | broken  | sound |
| <b>singular</b>          | 3%     | 99%   | 24%     | 86%   | 39%     | 67%   |
| <b>plural</b>            | 61%    | 97%   | 89%     | 98%   | 64%     | 87%   |
| <b>singular - plural</b> | 47%    | 94%   | 97%     | 68%   | 76%     | 56%   |

**Table 3:** 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

# Modeling Maltese Plurals

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

**Table 4:** 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

# Modeling Maltese Plurals

## Summary

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

# Modeling Maltese Plurals

Things to think about...

- we simplified Maltese considerably and modeled only two outcomes (sound and broken) BUT Maltese is more complicated
- 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?

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

# Modeling Maltese Plurals

## NDL Models 3 most frequent - Preparation

- three most frequent sound plural suffixes: -i (41%) , -iet (23%) and -ijiet (22%)
- three most frequent broken plural patterns: CCVVCVC (33%), (C)CVCVC (23%), and CCVVC (19%)
- outcomes: three most frequent suffixes and patterns + all other sound plurals combined as category *sound* and all other broken plurals combined as category *broken*
- paradigm coded as diphones provided the best predictions with binary outcome *broken* vs. *sound* → focus on diphones!

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



# Modeling Maltese Plurals

NDL Models 8 plural types - with the paradigm (singulars and plurals) as diphone cues

|               | broken A | broken B        | broken C       | broken         | sound iet       | sound ijiet     | sound i          | sound    |
|---------------|----------|-----------------|----------------|----------------|-----------------|-----------------|------------------|----------|
| broken A      | 6 (27%)  | 4 (18%)         | 7 (32%)        | 1 (5%)         | 2 (9%)          | 0 (0%)          | 0 (0%)           | 2 (9%)   |
| broken B      | 1 (5%)   | <b>14 (70%)</b> | 0 (0%)         | 4 (20%)        | 0 (0%)          | 0 (0%)          | 1 (5%)           | 0 (0%)   |
| broken C      | 1 (8%)   | 3 (23%)         | <b>7 (54%)</b> | 2 (15%)        | 0 (0%)          | 0 (0%)          | 0 (0%)           | 0 (0%)   |
| broken (rest) | 1 (11%)  | 0 (0%)          | 0 (0%)         | <b>6 (67%)</b> | 0 (0%)          | 0 (0%)          | 0 (0%)           | 2 (22%)  |
| sound iet     | 1 (3%)   | 1 (3%)          | 4 (10%)        | 6 (15%)        | <b>26 (67%)</b> | 1 (3%)          | 0 (0%)           | 0 (0%)   |
| sound ijiet   | 0 (0%)   | 0 (0%)          | 0 (0%)         | 0 (0%)         | 1 (2%)          | <b>44 (96%)</b> | 1 (2%)           | 0 (0%)   |
| sound i       | 0 (0%)   | 1 (1%)          | 2 (1%)         | 4 (3%)         | 0 (0%)          | 1 (1%)          | <b>127 (94%)</b> | 0 (0%)   |
| sound (rest)  | 2 (6%)   | 0 (0%)          | 5 (15%)        | 3 (9%)         | 1 (3%)          | 0 (0%)          | 6 (18%)          | 17 (50%) |

Table 5: NDL model with paradigm coded as diphones as cues

# Modeling Maltese Plurals

NDL Models 8 plural types - with the paradigm (singulars and plurals) as diphone cues

|               | broken A | broken B        | broken C       | broken         | sound iet | sound ijiet | sound i   | sound    |
|---------------|----------|-----------------|----------------|----------------|-----------|-------------|-----------|----------|
| broken A      | 6 (27%)  | 4 (18%)         | 7 (32%)        | 1 (5%)         | 2 (9%)    | 0 (0%)      | 0 (0%)    | 2 (9%)   |
| broken B      | 1 (5%)   | <b>14 (70%)</b> | 0 (0%)         | 4 (20%)        | 0 (0%)    | 0 (0%)      | 1 (5%)    | 0 (0%)   |
| broken C      | 1 (8%)   | 3 (23%)         | <b>7 (54%)</b> | 2 (15%)        | 0 (0%)    | 0 (0%)      | 0 (0%)    | 0 (0%)   |
| broken (rest) | 1 (11%)  | 0 (0%)          | 0 (0%)         | <b>6 (67%)</b> | 0 (0%)    | 0 (0%)      | 0 (0%)    | 2 (22%)  |
| sound iet     | 1 (3%)   | 1 (3%)          | 4 (10%)        | 0 (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 6: NDL model with paradigm coded as diphones as cues

# Modeling Maltese Plurals

NDL Models 8 plural types - with the paradigm (singulars and plurals) as diphone cues

|               | broken A | broken B        | broken C       | broken         | sound iet | sound ijiet | sound i   | sound    |
|---------------|----------|-----------------|----------------|----------------|-----------|-------------|-----------|----------|
| broken A      | 6 (27%)  | 4 (18%)         | 7 (32%)        | 1 (5%)         | 2 (9%)    | 0 (0%)      | 0 (0%)    | 2 (9%)   |
| broken B      | 1 (5%)   | <b>14 (70%)</b> | 0 (0%)         | 4 (20%)        | 0 (0%)    | 0 (0%)      | 1 (5%)    | 0 (0%)   |
| broken C      | 1 (8%)   | 3 (23%)         | <b>7 (54%)</b> | 2 (15%)        | 0 (0%)    | 0 (0%)      | 0 (0%)    | 0 (0%)   |
| broken (rest) | 1 (11%)  | 0 (0%)          | 0 (0%)         | <b>6 (67%)</b> | 0 (0%)    | 0 (0%)      | 0 (0%)    | 2 (22%)  |
| sound iet     | 1 (3%)   | 1 (3%)          | 4 (10%)        | 0 (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 7: NDL model with paradigm coded as diphones as cues

# Modeling Maltese Plurals

NDL Models 8 plural types - with the paradigm (singulars and plurals) as diphone cues

|               | broken A | broken B        | broken C       | broken         | sound iet | sound ijiet | sound i   | sound    |
|---------------|----------|-----------------|----------------|----------------|-----------|-------------|-----------|----------|
| broken A      | 6 (27%)  | 4 (18%)         | 7 (32%)        | 1 (5%)         | 2 (9%)    | 0 (0%)      | 0 (0%)    | 2 (9%)   |
| broken B      | 1 (5%)   | <b>14 (70%)</b> | 0 (0%)         | 4 (20%)        | 0 (0%)    | 0 (0%)      | 1 (5%)    | 0 (0%)   |
| broken C      | 1 (8%)   | 3 (23%)         | <b>7 (54%)</b> | 2 (15%)        | 0 (0%)    | 0 (0%)      | 0 (0%)    | 0 (0%)   |
| broken (rest) | 1 (11%)  | 0 (0%)          | 0 (0%)         | <b>6 (67%)</b> | 0 (0%)    | 0 (0%)      | 0 (0%)    | 2 (22%)  |
| sound iet     | 1 (3%)   | 1 (3%)          | 4 (10%)        | 0 (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 8: NDL model with paradigm coded as diphones as cues

- broken plurals are often confused with other broken plural types

# Modeling Maltese Plurals

NDL Models 8 plural types - with the paradigm (singulars and plurals) as diphone cues

|               | broken A | broken B | broken C | broken  | sound iet       | sound ijiet     | sound i          | sound    |
|---------------|----------|----------|----------|---------|-----------------|-----------------|------------------|----------|
| broken A      | 6 (27%)  | 4 (18%)  | 7 (32%)  | 1 (5%)  | 2 (9%)          | 0 (0%)          | 0 (0%)           | 2 (9%)   |
| broken B      | 1 (5%)   | 14 (70%) | 0 (0%)   | 4 (20%) | 0 (0%)          | 0 (0%)          | 1 (5%)           | 0 (0%)   |
| broken C      | 1 (8%)   | 3 (23%)  | 7 (54%)  | 2 (15%) | 0 (0%)          | 0 (0%)          | 0 (0%)           | 0 (0%)   |
| broken (rest) | 1 (11%)  | 0 (0%)   | 0 (0%)   | 6 (67%) | 0 (0%)          | 0 (0%)          | 0 (0%)           | 2 (22%)  |
| sound iet     | 1 (3%)   | 1 (3%)   | 4 (10%)  | 6 (15%) | <b>26 (67%)</b> | 1 (3%)          | 0 (0%)           | 0 (0%)   |
| sound ijiet   | 0 (0%)   | 0 (0%)   | 0 (0%)   | 0 (0%)  | 1 (2%)          | <b>44 (96%)</b> | 1 (2%)           | 0 (0%)   |
| sound i       | 0 (0%)   | 1 (1%)   | 2 (1%)   | 4 (3%)  | 0 (0%)          | 1 (1%)          | <b>127 (94%)</b> | 0 (0%)   |
| sound (rest)  | 2 (6%)   | 0 (0%)   | 5 (15%)  | 3 (9%)  | 1 (3%)          | 0 (0%)          | 6 (18%)          | 17 (50%) |

Table 9: NDL model with paradigm coded as diphones as cues

# Modeling Maltese Plurals

NDL Models 8 plural types - with the paradigm (singulars and plurals) as diphone cues

|               | broken A | broken B | broken C | broken  | sound iet       | sound ijiet     | sound i          | sound    |
|---------------|----------|----------|----------|---------|-----------------|-----------------|------------------|----------|
| broken A      | 6 (27%)  | 4 (18%)  | 7 (32%)  | 1 (5%)  | 2 (9%)          | 0 (0%)          | 0 (0%)           | 2 (9%)   |
| broken B      | 1 (5%)   | 14 (70%) | 0 (0%)   | 4 (20%) | 0 (0%)          | 0 (0%)          | 1 (5%)           | 0 (0%)   |
| broken C      | 1 (8%)   | 3 (23%)  | 7 (54%)  | 2 (15%) | 0 (0%)          | 0 (0%)          | 0 (0%)           | 0 (0%)   |
| broken (rest) | 1 (11%)  | 0 (0%)   | 0 (0%)   | 6 (67%) | 0 (0%)          | 0 (0%)          | 0 (0%)           | 2 (22%)  |
| sound iet     | 1 (3%)   | 1 (3%)   | 4 (10%)  | 6 (15%) | <b>26 (67%)</b> | 1 (3%)          | 0 (0%)           | 0 (0%)   |
| sound ijiet   | 0 (0%)   | 0 (0%)   | 0 (0%)   | 0 (0%)  | 1 (2%)          | <b>44 (96%)</b> | 1 (2%)           | 0 (0%)   |
| sound i       | 0 (0%)   | 1 (1%)   | 2 (1%)   | 4 (3%)  | 0 (0%)          | 1 (1%)          | <b>127 (94%)</b> | 0 (0%)   |
| sound (rest)  | 2 (6%)   | 0 (0%)   | 5 (15%)  | 3 (9%)  | 1 (3%)          | 0 (0%)          | 6 (18%)          | 17 (50%) |

Table 10: NDL model with paradigm coded as diphones as cues

# Modeling Maltese Plurals

NDL Models 8 plural types - with the paradigm (singulars and plurals) as diphone cues

|               | broken A | broken B | broken C | broken  | sound iet | sound ijiet | sound i   | sound    |
|---------------|----------|----------|----------|---------|-----------|-------------|-----------|----------|
| broken A      | 6 (27%)  | 4 (18%)  | 7 (32%)  | 1 (5%)  | 2 (9%)    | 0 (0%)      | 0 (0%)    | 2 (9%)   |
| broken B      | 1 (5%)   | 14 (70%) | 0 (0%)   | 4 (20%) | 0 (0%)    | 0 (0%)      | 1 (5%)    | 0 (0%)   |
| broken C      | 1 (8%)   | 3 (23%)  | 7 (54%)  | 2 (15%) | 0 (0%)    | 0 (0%)      | 0 (0%)    | 0 (0%)   |
| broken (rest) | 1 (11%)  | 0 (0%)   | 0 (0%)   | 6 (67%) | 0 (0%)    | 0 (0%)      | 0 (0%)    | 2 (22%)  |
| 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 11: NDL model with paradigm coded as diphones as cues

- sound plurals are confused with other sound plural types and sometimes with other broken plural types

## Discussion & Conclusion

- we used different NDL models to predict Maltese sound and broken plurals
- cues: singulars, plurals, paradigm coded as single phones, diphones or triphones
- outcomes: broken vs. sound (1st set of models) or 3 most frequent sound suffixes and broken plural patterns (2nd model)
- to correctly predict Maltese sound and broken plurals, NDL needs information about the plurals.
- our results are in line with the Word and Paradigm model of morphological processing

(Blevins, 2016)



Grazzi ħafna!

## References I

- Arppe, A., Hendrix, P., Milin, P., Baayen, H., Sering, T., & Shaoul, C. (2015). *Ndl: Naive Discriminative Learning*. <https://CRAN.R-project.org/package=ndl>.
- Baayen, H. R. (2011). Corpus linguistics and naive discriminative learning. *Brazilian Journal of Applied Linguistics*, 11, 295–328.
- Blevins, J. (2016). *Word and Paradigm Morphology*. Oxford, Oxford University Press.
- Danks, D. (2003). Equilibria of the Rescorla-Wagner model. *Journal of Mathematical Psychology*, 47, 109–121.
- Evert, S., & Arppe, A. (2015). Some theoretical and experimental observations on naive discriminative learning. In *Proceedings of the 6th Conference on Quantitative Investigations in Theoretical Linguistics (qitl-6)*, Tübingen, Germany.
- Nieder, J., van de Vijver, R., & Mitterer, H. (2020a). Knowledge of Maltese singular-plural mappings. *Morphology*.

- Nieder, J., van de Vijver, R., & Mitterer, H. (2020b). *Priming Maltese Plurals: Representation of sound and broken plurals in the mental lexicon* [Manuscript submitted for publication]. Manuscript submitted for publication.
- Plag, I., & Balling, L. W. (2016). Derivational morphology: An integrative perspective on some fundamental issues (V. Pirelli, I. Plag, & W. U. Dressler, Eds.). In V. Pirelli, I. Plag, & W. U. Dressler (Eds.), *Word knowledge and word usage: A cross-disciplinary guide to the mental lexicon*. Berlin, New York, De Gruyter.
- R Core Team. (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.
- Ramscar, M., Dye, M., & McCauley, S. (2013). Error and expectation in language learning: The curious absence of 'mouses' in adult speech. *Language*, 89(4), 760–793.

- Ramscar, M., Yarlett, D., Dye, M., Denny, K., & Thorpe, K. (2010). The Effects of Feature-Label-Order and Their Implications for Symbolic Learning. *Cognitive Science*, 34(6), 909–957.  
<https://doi.org/10.1111/j.1551-6709.2009.01092.x>
- Rescorla, R., & Wagner, A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement (A. Black & W. Prokasy, Eds.). In A. Black & W. Prokasy (Eds.), *Classical conditioning ii: Current research and theory*. New York, Appleton-Century-Crofts.
- Zwitserslood, P. (2018). Processing and representation of morphological complexity in native language comprehension and production (G. Booij, Ed.). In G. Booij (Ed.), *The construction of words: Advances in Construction Morphology*. Berlin: Springer.

# Acknowledgements

This research was funded by the Deutsche Forschungsgemeinschaft (FOR2373, Projects VI 223/3-1 and VI 223/3-2), which we gratefully acknowledge.

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.