

Productivity of the broken plural in Maltese

In this research, we examine the productivity of the broken plural in Maltese. Using machine learning techniques and behavioral methods, we show that the broken plural is able to be predicted by a singular noun's CV structure.

Using a logistic regression classifier, we are able to correctly predict the plural CV structure of the singular noun with 69% accuracy when classifying among all available options and 80% accuracy when classifications are restricted by the structure of the singular (32% and 4% respective relative increases over a strong generalized context model (Dawdy-Hesterberg & Pierrehumbert, 2014) baseline).

We are further investigating the results obtained from the machine learning using behavioral methods. In a Wug Test (following Berko, 1958), Maltese native speaker participants are shown nonsense words conforming to the CV structures found above, randomly paired with an imaginary animal picture (from Ohala, 1999). Each participant provides a response to the carrier phrase "This is a [wug]. If you saw three of them, you would say, 'There are three [wugs].'"

Using these results together, we will be able to determine whether humans are sensitive to the same features that were useful to the machine learning algorithm. Previous research suggests that humans are sensitive to CV structures when asked to form novel words (e.g., Galea, 2011), and we will expand on those results in Maltese broken plurals. Research showing that some participants are sensitive to the differences in words that sound Semitic vs. those that sound Indo-European (e.g., Twist, 2006; Drake, in prep) may also come into play here; however, it can be seen that broken plurals can be applied to non-Semitic loanwords, such as *gakketta* 'jacket' ~ *gkieket* 'jackets'. This suggests that using a broken plural productively may not be confined to the Semitic sub-lexicon.