

“All mimsy were the borogoves” – A discriminative learning model of morphological knowledge in pseudo-word inflection

Jessica Nieder^{a*}, Ruben van de Vijver^a and Fabian Tomaschek^b

^aDepartment of General Linguistics, Heinrich-Heine-Universität Düsseldorf, Germany;

^bDepartment of General Linguistics, Eberhard-Karls-Universität Tübingen, Germany

Abstract

Grammatical knowledge of native speakers has often been investigated in so-called *wug* tests, in which participants have to inflect pseudo-word forms (wugs). Typically it has been argued that in inflecting these pseudo-words, speakers apply their knowledge of word formation processes. However, it remains unclear what exactly this knowledge is and how it is learned.

According to one theory, the knowledge is best characterized as abstractions and rules that specify how units can be combined. Another theory maintains that it is best characterized by analogy. In both cases the knowledge is learned by association based on positive evidence alone.

In this paper, we model the classification of pseudo-words to Maltese plural classes on the basis of phonetic input using a shallow neural network trained with an error-driven learning algorithm. We demonstrate that the classification patterns mirror those of Maltese native speakers in a wug test. Our results indicate that speakers rely on gradient knowledge of a relation between the phonetics of whole words and plural classes, which is learned in an error-driven way.

Keywords Discriminative Learning; Naive Discriminative Learning; Wug Test; Computational Modelling; Maltese Plurals; Production

*Email: nieder@phil.hhu.de

1 Introduction

As the first part of our title – the end of Lewis Carroll’s ‘Jabberwocky’ – demonstrates, native speakers are able to identify the type of words these nonces represent and inflect the pseudo-words according to the morphological rules in their native language. Wug tests are used to exploit this ability. On the basis of the data gathered in these tests, researchers inspect morphologically complex (pseudo)words produced by speakers and draw, on the basis of their characteristics, conclusions about the nature of their morphological knowledge (Berko, 1958). This raises the question what exactly it is that speakers know about morphology and how this knowledge is learned. The present study investigates the nature of the speakers’ knowledge in wug tests. To this end we modeled the results of a wug-experiment on Maltese performed by Nieder, van de Vijver, et al. (2021a). We did so in a simple two-layer fully connected neural network. We tested the performance of the network by correlating the proportions of answers given by the participants in the wug-experiment, and the proportions of plural classes from a Maltese corpus with the proportions of answers predicted by the network.

To anticipate our results, we find high correlations between the participants’ and the network’s performance and between the plural classes in the corpus and the network’s performance. We conclude that the knowledge of native speakers is best characterized as consisting of generalizations that are based on whole words, and that are learned in an error-driven fashion.

1.1 Theoretical perspectives on the wug test

Ever since Berko concluded, based on her famous wug test, that English children know the morphology of their native language because they were able to inflect novel words grammatically during a production experiment, the question is raised what exactly it is that speakers know about morphology. Much work has been devoted to answer this question. In the following subsections we discuss the answers presented from three perspectives: abstraction-based approaches in section 1.1.1, analogy-based approaches in section 1.1.2, and approaches based on error-driven learning in section 1.1.3.

1.1.1 Abstraction-based approaches

In the generative conceptualization of language, knowledge of language is represented in the form of abstract statements that map a form stored in memory onto its surface form. These statements are either formalized as rules (Albright & Hayes, 2003; Chomsky, 1957; Chomsky, 1965; Chomsky

& Halle, 1968; Pinker, 1989), or as constraints (Prince & Smolensky, 2004). Rules are explicit statements that describe a context and a change of sounds applied when the context is met. An important question is how such rules can be learned. Albright and Hayes (2003) propose that they are induced by incrementally comparing word forms. In the case of learning past tense forms, the comparison is between stems and past tense forms, for example *tak* ‘talk’ and *takt* ‘talked’. This pair is construed as a word-specific rule as in (1). The left part of the rule is called the change and the right part of the rule is called the context. This rule is to be read as follows: an empty category (\emptyset) is changed to [t] in the context of the word [tak] if it is indexed with the morphological feature [+past].

$$\emptyset \longrightarrow t/[tak______]_{+past} \quad (1)$$

Going through further pairs in the lexicon, the pair *wak* ‘walk’ and *wakt* ‘walked’ is found. This, again, is used to set up a word specific rule (2).

$$\emptyset \longrightarrow t/[wak______]_{+past} \quad (2)$$

These rules can be made more general, by keeping the similar parts and trying to describe the differing parts as a natural class or a variable. The more general rule that results from this process is given in 3. It has replaced the initial segments of the verbs with their natural class [-syllabic].

$$\emptyset \longrightarrow t/[\text{-syllabic}] [\text{-ak}]______]_{+past} \quad (3)$$

Rules are then given a score, by counting how often they are applied and dividing it by the number of words in which the context of a rule is found in a corpus. This proportion is then adjusted to ensure that rules which contain contexts that occur less often are given less weight than rules that apply in contexts that occur more often. For example, if speakers, who have induced such rules, encounter a pseudo-word, e.g. *mak*, and are asked to provide its past tense, they would use the general rule to create *makt*.

Albright and Hayes used this method to model the knowledge of English native speakers of the past tense of English verbs. They used the English past tense to argue that the rules that the model induces are indeed used by native speakers of English, when they are asked to provide the past tense for pseudo verbs. To this aim, they provided the model with stems and past tense forms, which allowed it to deduce rules. The model’s rules were then used to provide the past tense for pseudo verbs. The inflected pseudo-words were finally used as items in two experiments. In the first, the participants were asked to provide past tense forms for pseudo stems predicted by their

model. In the second experiment participants were asked to rate the well-formedness of the past tense forms produced by their model. The correlation between the rating of regular verbs (the ones created from rules that add a regular allomorph) and the scores of the model was $\rho = 0.714$, and the correlation between the rating of irregular verbs and the scores of the model was $\rho = 0.480$. This shows that the predictions of the model correlate with the behavior of native speakers, and, therefore, the model was interpreted to be a good model of the knowledge of native speakers about their past tense system. Similar results have been reported for other languages, e.g. Italian (Albright, 2002b), Lakotha (Albright, 2002a) and Spanish (Albright et al., 2001).

Yet, despite its success, Albright and Hayes’s *Minimal Generalization Learner* cannot serve as a general model of knowledge of morphology. The reason is that it is built upon rules that affect segments in their immediate contexts. However, there are a number of languages in which grammatical functions are expressed by prosodic changes rather than by contextual segmental changes. This is the case in the Semitic language Maltese which will be the target of the present study. In Maltese, prosodic changes such as in the noun plural inflection [fardal] ‘apron’ \rightarrow [fra:ɖal] ‘aprons’ (Nieder, van de Vijver, et al., 2021a) cannot be captured by linear rules.

We are by no means the first to observe this: The realization that some processes cannot be captured by linear rules led to the development of autosegmental phonology (Goldsmith, 1979; Leben, 1973; Woo, 1969). It is difficult to adapt the *Minimal Generalization Learner* to incorporate the ideas from autosegmental phonology. The rules would have to be able to refer to syllable positions, and that is something rules in this form simply can not achieve. As a remedy, the rules could make use of templates which specify CV-templates, that is sequences of consonants and vowels the words are made up of (Marantz, 1982). However, as McCarthy and Prince (1990) argue, such templates are arbitrary and need to be defined in terms of prosodic units, such as morae, syllables or feet.

In the examples from Maltese given above (sg. [fardal] \rightarrow pl. [fra:ɖal]), the first syllable of the singular *far* needs to be mapped onto the first syllable of the plural *fra:*. For this system to be general, in the sense that it can accommodate all language types, it is not clear to what extent *fra:* would be a basic prosodic unit, since it has a complex onset and a long vowel – and accordingly is considered to be a marked syllable.

If we try to define the preferred plural in terms of the constraint interaction of Maltese, we also run into problems. Following Optimality Theory (Prince & Smolensky, 2004), it is impossible to pin down the complex syllable *fra:* as optimal, because it has a complex onset and there must also

be a candidate that has a simple onset. The candidate with the complex onset violates the constraint that requires segments that are contiguous in the input to be contiguous in the output, and this constraint is not violated by the candidate with the simple onset. It is therefore not clear by what general mechanism rules or constraints that map a singular onto a plural in Maltese could be induced.

Another class of rule-based models is formed by speech production models that conceive of speech production as a modular process. During this process, semantic concepts select an appropriate lemma (e.g. ‘walk’ for a movement behavior), which in turn activates syntactic and morphological properties of the lemma, and these, in turn, activate its phonological and phonetic properties. The *Utterance Generator* (Fromkin, 1971, 1988) and *The Theory of Lexical Access* by Levelt and colleagues (Indefrey & Levelt, 2000; Levelt et al., 1999; Roelofs & Ferreira, 2019) belong in this class. An overview of such models of speech production can be found in (Tucker & Tomaschek, to appear).

Of interest for the interpretation of wug tests are, within the framework of speech production models, the processes at the morphemic stage. At this stage, the lemma selects the appropriate morphemes which encode the intended meaning, e.g. the stem for ‘walk’ and s for ‘second person singular present tense’. Once concatenated, they are passed to phonological encoding resulting in the phonological form, e.g. waks. Phonological sequences like this one activate syllable representations which drive articulation.

This perspective is implemented in computational models of speech production, such as *Weaver++* (Roelofs, 1997) or the *Spreading-Activation Model* by Dell (1986). Within this perspective, when speakers have to inflect a new word, as is the case in a pseudo-word test, it is these rule-based processes that come into play. The system selects the appropriate morphemes, concatenates them and, as a result, speakers pronounce the plural ‘wugs’ for the noun ‘wug’ or the past tense ‘wugged’ for the verb ‘to wug’. However, similarly to the rules of the previously described *Minimal Generalization Learner*, it is unlikely that the embedded linear rules are able to handle the morphologically non-linear patterns of Maltese.

1.1.2 Analogy-based approaches

In another group of models, morphological knowledge emerges from stored – inflected – word forms. Generalizations are formed on the basis of similarities among these stored forms and this knowledge is applied when a form needs to be inflected. Rumelhart and McClelland (1986) proposed an early, computationally implemented analogical model in the form of a simple recurring

neural network, with which the formation of English past tense verbs on the basis of present tense verbs was modeled. They argued that the formation of morphologically complex word forms does not rely on the application of rules, but on analogies. These analogies are formed by learning the associations between the phonetic similarity of word forms, on the one hand, and their grammatical class, on the other hand. The model is trained to produce past tense forms on the basis of present tense forms. After training, the model is asked to produce a past tense form of words it has not seen yet. It does so by assessing the phonetic similarity of the new word to words it has been trained with, and producing past tense forms that closely resemble forms in phonetically similar word forms from the training. Other analogy-based theories have been proposed in Bybee (2010), Daelemans and Van den Bosch (2005) and Skousen (1989).

Similarity in computational analogical models is assessed by means of k NN algorithms. This is for example the case in the *Tilburg Memory-Based Learner* (TiMBL), which is a computationally implemented memory-based analogical model proposed by Daelemans and Van den Bosch (2005). Similarity among forms is assessed by comparing for each position in the words how similar the sounds in that position are. The similarity can be computed in a strict way, in which a p resembles a p and nothing else, or in a more loose way based on phonetic features in which a p resembles a p , but also a b and an f (Daelemans & Van den Bosch, 2005) in certain positions. The model proposed by Skousen (1989) works in a similar way, but has a different computational implementation than TiMBL.

Despite TiMBL’s success in classifying Maltese nouns, as reported in Nieder, Tomaschek, et al. (2021), there are two issues with a memory-based analogical model to capture the knowledge of the Maltese noun plural system. The first issue is that the similarity of words is assessed by comparing words in an edge-aligned fashion. The researcher decides whether the model aligns all words from the left or the right side of the word.

Yet speakers have nobody telling them where the comparison between words has to start, or where it has to finish. Unfortunately, phonological similarity among words without pre-specifying where to start the comparison is currently beyond the reach of analogical models.

Another issue is that analogical models capture similarity by finding the largest groups of forms that are most phonologically similar to a form for which a generalization needs to be made (Daelemans & Van den Bosch, 2005; Keuleers et al., 2007). While mathematically very powerful, learning only on the basis of similarity disregards that human learning is best modeled as an error-driven, discriminative process in which the goal is to minimize prediction errors (Ramscar, Dye, & Klein, 2013; Ramscar et al., 2010).

1.1.3 Discriminative approaches

In discriminative approaches morphological knowledge emerges from the way in which the sounds of a word are used to discriminate its meaning from among other meanings. In contrast to the associative perspective on word learning (Ellis, 2006a, 2006b), the discriminative strength of the sounds of a word is learned in an error-driven fashion, by minimizing prediction errors. This goal is accomplished by taking into account occurrences and non-occurrences between events, as well as the established learning history. As Ramscar, Dye, and McCauley (2013), Ramscar and Yarlett (2007), and Ramscar et al. (2010) have repeatedly demonstrated, this process can be modelled by using an error-driven learning rule, as proposed by Rescorla and Wagner (1972a).

To give a short example of how discrimination and error-driven learning work, consider the German verb forms *schreibe* “(I) write, am writing” and *schreiben* “(we, they) write, are writing”. In this example we will use syllables, indicated by a period as cognitively plausible units. Upon hearing the syllables in *schrei.be* the learner increases the weights on the association between the syllables of the verb and its meaning. Upon hearing *schrei.ben* the weights of *schrei* and *ben* to its meaning are increased, and the weight of *schrei* to the meaning of “(I) write, am writing” is decreased. This is because the *schrei* predicts “(I) write, am writing” (on the basis of previous experience), but this prediction is wrong for the verb form *schrei.ben* (Ramscar et al., 2010; Rescorla, 1988). It is important to emphasize that morphology can emerge from discriminative learning, even though there is no morphological structure in the input because syllable structure and morphological structure are not isomorph.

Discriminative Learning, as proposed by Ramscar, Dye, and McCauley (2013), Ramscar and Yarlett (2007), and Ramscar et al. (2010) has been applied to a great deal of linguistic phenomena: morphological learning (Ramscar et al., 2010), word recognition (Arnold et al., 2017; Shafaei-Bajestan et al., 2020), non-word processing (Chuang et al., 2019), phonetic learning (Nixon & Tomaschek, 2021), morpho-phonetic interactions (Schmitz et al., 2021; Tomaschek et al., 2019), classification of morphologically complex words (Baayen, 2011; Nieder, Tomaschek, et al., 2021), priming (Baayen & Smolka, 2020), induction of morphological knowledge (Divjak et al., 2020), modelling frequency effects (Baayen et al., 2013), lexical decision (Milin et al., 2017), and compounding (Pham & Baayen, 2015).

The theory of Discriminative Learning is computationally implemented as a simple two-layer input-output neural network in the *Naive Discriminative Learner* (NDL, Arppe et al. (2018) – see the Appendix for the mathematical

details of the two-layer network and the way in which the weights of the association between cues and outcomes are calculated). NDL is moreover embedded within the theory of the Discriminative Lexicon (Baayen et al., 2019).

By modelling the results of a wug test for a language with an intricate plural system, we bring this theory to bear on an even wider range of phenomena. This brings us to the aims and scopes of the present study.

1.2 The present study

In NDL, we model the results of a production experiment in Maltese by Nieder, van de Vijver, et al. (2021a), who studied how native speakers of Maltese pluralize words and pseudo-words in a wug test. In this way, we address how the morphology of a language is represented by its speakers and how this knowledge is applied when inflecting novel words. The result of our modelling approach is a new perspective on morphological knowledge for languages with linear and prosodic (non-linear) morphological processes, not in the form of rules or analogy, but as the result of discriminative learning.

In the next section, we will describe the phonological and morphological characteristics of Maltese, before focusing on Nieder, van de Vijver, et al.’s (2021a) wug experiment. After describing our modelling approach, we present our modelling results. We conclude the paper with a discussion in which we relate the present findings to theories of morphological knowledge.

2 Maltese

Having discussed the theoretical background of the present study, we next present the material and experimental background for the current computational study.

2.1 Broken and sound plurals

Maltese is a non-concatenative Semitic language that developed from a Maghrebi Arabic variety and is highly influenced by several concatenative languages (e.g. English, Sicilian, Italian) due to the colonial history of the Maltese islands. As a consequence, the Maltese noun plural system shows an opposition of concatenative *sound* plurals, e.g. *annimal* - *annimali* ‘animals’, vs. non-concatenative *broken* plurals, e.g. *qattus* - *qtates* ‘cats’.

To express a sound plural, Maltese native speakers have to select one of 12 different suffixes (Nieder, van de Vijver, et al., 2021a). In the case

of broken plurals the choice is based on 11 different broken plural patterns (Nieder, van de Vijver, et al., 2021a; Schembri, 2012). Table 1 taken from Nieder, Tomaschek, et al. (2021) displays all Maltese sound plural suffixes and broken plural patterns. This variety clearly raises the question about the grammatical knowledge that emerges from this kind of system and thus, what kind of information native speakers use to inflect novel word forms.

Singular	Plural	Gloss	Plural Class	Proportions
<i>karta</i>	<i>karti</i>	‘paper’	sound, <i>-i</i>	1242 words
<i>omm</i>	<i>ommiġiet</i>	‘mother’	sound, <i>-iġiet</i>	430 words
<i>rixa</i>	<i>rixiet</i>	‘feather’	sound, <i>-iet</i>	409 words
<i>giddieb</i>	<i>giddieba</i>	‘liar’	sound, <i>-a</i>	109 words
<i>mehlus</i>	<i>mehlusin</i>	‘freed’	sound, <i>-in</i>	91 words
<i>kuxin</i>	<i>kuxins</i>	‘cushion’	sound, <i>-s</i>	53 words
<i>triq</i>	<i>triqat</i>	‘street’	sound, <i>-at</i>	51 words
<i>sid</i>	<i>sidien</i>	‘owner’	sound, <i>-ien</i>	5 words
<i>bahri</i>	<i>bahrin</i>	‘sailor’	sound, <i>-n</i>	2 words
<i>hati</i>	<i>hatjin</i>	‘guilty’	sound, <i>-jin</i>	1 words
<i>spalla</i>	<i>spallejn</i>	‘shoulder’	sound, <i>-ejn/ajn</i>	1 word
<i>sieq</i>	<i>saqajn</i>	‘foot’	sound, <i>-ejn/ajn</i>	1 word
<i>qiegh</i>	<i>qighan</i>	‘bottom’	sound, <i>-an</i>	1 word
<i>fardal</i>	<i>fradal</i>	‘aprons’	broken, CCVVCVC	258 words
<i>birra</i>	<i>birer</i>	‘beers’	broken, (C)CVCVC	181 words
<i>kbir</i>	<i>kbar</i>	‘big (pl.)’	broken, CCVVC	148 words
<i>bitħa</i>	<i>btieħi</i>	‘yards’	broken, CCVVCV	71 words
<i>ftira</i>	<i>ftajjar</i>	‘type of bread (pl.)’	broken, CCVjjVC	64 words
<i>sider</i>	<i>isdra</i>	‘chests’	broken, VCCCV	31 words
<i>marid</i>	<i>morda</i>	‘sick persons’	broken, CVCCV	14 words
<i>ghodda</i>	<i>ghodod</i>	‘tools’	broken, (gh)VCVC	0 words
<i>elf</i>	<i>eluf</i>	‘thousands’	broken, VCVC	5 words
<i>gharef</i>	<i>ghorrief</i>	‘wise men’	broken, CVCCVVC(V)	6 words
<i>ghama</i>	<i>ghomja</i>	‘blind persons’	broken, (gh)VCCV	1 word

Table 1: Maltese broken and sound plurals, table taken from Nieder, Tomaschek, et al. (2021). The proportions in the rightmost column are based on the data set that was used for this study. The pattern *(gh)VCVC* identified by Schembri (2012) did not have any instances in Nieder, Tomaschek, et al.’s (2021) and in our data set.

2.2 Production study on Maltese plurals

The current study intends to predict the results of the wug experiment on Maltese performed by Nieder, van de Vijver, et al. (2021a) in which they tested the mapping of Maltese singulars onto plurals. Concretely, they investigated to what degree the phonological form of the (pseudo) singular determined the plural class of the plural form (broken / sound).

Eighty adult native speakers of Maltese were asked to pluralize frequent and infrequent existing Maltese singulars as well as phonotactically legal pseudo singulars. Pseudo-words were constructed by changing the consonants, the vowels or both, consonants and vowels, of existing words. Following this procedure the phonotactics of Maltese word forms was maintained while creating non-existing singulars that show different degrees of similarity to real Maltese words.

Based on a wug test setup, Nieder, van de Vijver, et al. (2021a) presented the experimental items with pictures of existing things (existing words) and fantasy animals (pseudo-words). The primary research question was to investigate if Maltese native speakers make use of a morphological default rule as a productive pluralization strategy as proposed by Marcus, Brinkmann, et al. (1995). In this case, some sound plural suffixes or broken plural patterns were expected to be used more frequently than others regardless of their frequency in the lexicon. This hypothesis directly reflected the assumption of a dual-mechanism approach of morphological processing (Marcus, Brinkman, et al., 1995; Martinet, 1965; Pinker & Prince, 1991; Pinker & Prince, 1988). If Maltese speakers instead based their pluralization strategy on the similarity of the pseudo-words with existing words in the lexicon, Nieder, van de Vijver, et al. (2021a) took this as evidence for an analogical mechanism of morphological processing (Albright & Hayes, 2003; Daelemans & Van den Bosch, 2005; Skousen, 1992).

Nieder, van de Vijver, et al. (2021a) found a positive correlation between the proportions of sound suffixes and broken plural patterns in a Maltese corpus and the proportions of suffixes and plural patterns used for the pluralization of pseudo-words. Moreover, wrong plural answers for existing nouns mainly occurred in the case of infrequent broken plural nouns. This indicates that Maltese native speakers were more certain about sound plurals than about broken plurals.

As for the pseudo-words, the Maltese participants demonstrated a great amount of variation. Individual items were pluralized with both sound and broken patterns. This indicates that speakers of Maltese often accept several plural forms for one singular, even if they would only use one of these forms themselves. Overall, these results showed that speakers did not make use of a

default rule. Rather, they base their plural choice in analogy to similar forms and their frequency in their lexicon (Nieder, van de Vijver, et al., 2021a).

3 Methods

The following sections are dedicated to the computational approach presented in this study. We will first focus on the composition of the data sets used for training and testing the network before reporting details about the model design. Finally, we will present the results of our modelling efforts.

3.1 Data sets

In this study, we model the results of the production experiment by Nieder, van de Vijver, et al. (2021a) described in section 2.2. To be able to do so, two different data sets were used.

The first data set is a collection of 3174 singular-plural pairs (after the removal of unclear data). This data set was originally compiled by Nieder, van de Vijver, et al. (2021a) from the MLRS Korpus Malti 2.0 and 3.0 and a broken plural list collected by Schembri (2012). The compiled list of all nouns from the corpus was then matched with data from the only dictionary Ġabra (Camilleri, 2013) by using the free corpus tool *Coquery* (Kunter, 2017). It was already used by Nieder, Tomaschek, et al. (2021) for their modelling study and slightly reduced for this study due to unclear data. The data set contains a quasi-phonetic transcription of the singular and plural forms obtained by replacing each phone with exactly one letter or symbol (see appendix 5.2). Figure 1 displays the proportion of the different Maltese sound plural suffixes and broken plural patterns in the training data set. As can be seen in the Figure, sound plurals represent the majority of plural forms. This data set served as a training data set for our NDL models.

The second data set was taken from Nieder, van de Vijver, et al. (2021a) and contained 8972 observations from their production study. We added a quasi-phonetic transcription of the singular forms in the same way as was done for the training data set. This data set was used as a test data set for the NDL model.

Figure 2 demonstrates that the proportions of pluralization types in the experiment is very similar to the proportions in the Maltese lexicon (Figure 1), with the sound plural suffixes *-i*, *-ijiet* and *-iet* being used in most cases. As Nieder, van de Vijver, et al. (2021a) have established, their participants used the most frequent plural classes in the Maltese lexicon to pluralize words they have never heard before.

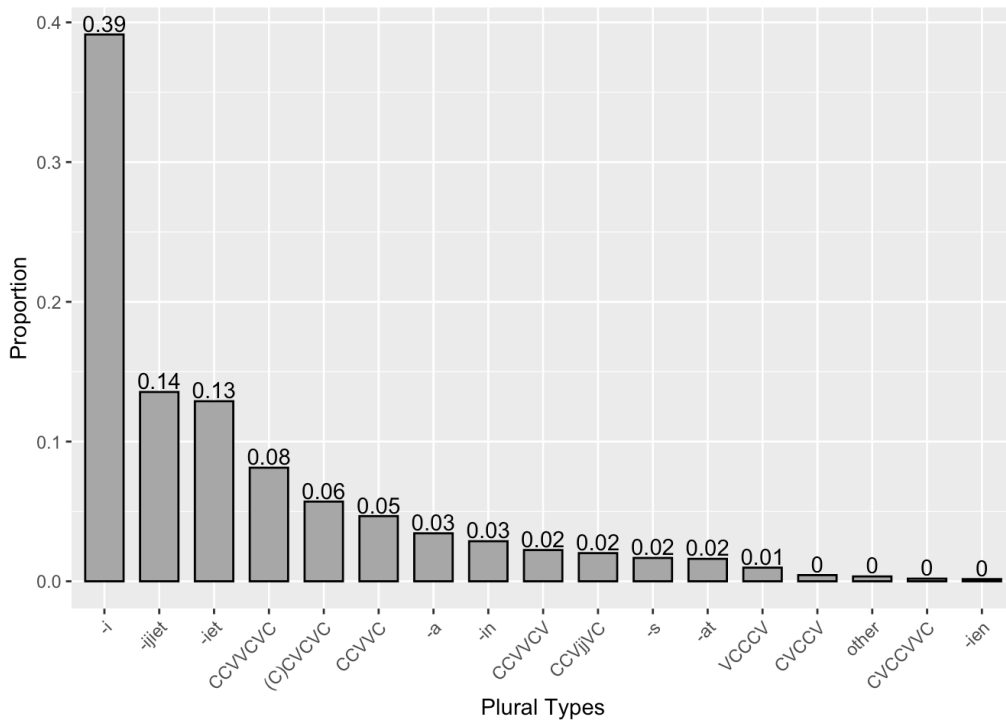


Figure 1: Proportion of sound (suffixes) and broken (CV structures) plural classes in the training data set. *Other* contains a collection of infrequent plurals with less than 5 instances from 7 different categories (3 broken plural classes and 4 sound plural classes) with a total number of 11 words.

Having discussed our material, we will describe our modelling approach as well as the design of our models that we used to investigate to what degree NDL is capable predicting the results of a wug test in the next section.

3.2 Modelling approach and model design

Modelling experiments in NDL are typically performed in a two step process. In the first step, a network is trained to associate input and output; in the second step, the trained network is used to predict the output on the basis of newly presented input.

In our case, in the first step we trained the NDL network to learn to associate provided word form cues with plural class outcomes. Training was done with the Danks equilibrium equation (Danks, 2003) that, in contrast to the error-driven learning equations by Rescorla and Wagner (Rescorla & Wagner, 1972a), allow fast computation of connection weights between cues and outcomes. Both equations can be found in the Appendix (Section 5.2).

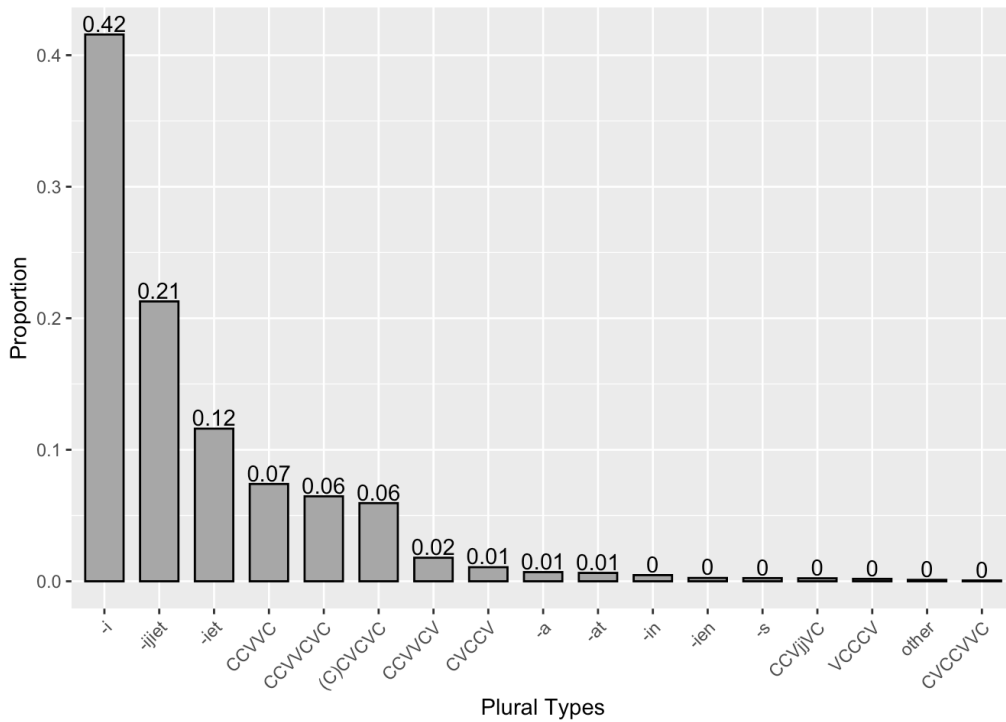


Figure 2: Proportion of plural classes in the experimental data set. The category *other* contains infrequent plurals with less than 5 words from 5 different categories (2 broken plural classes, 3 sound plural classes) with a total number of 14 words.

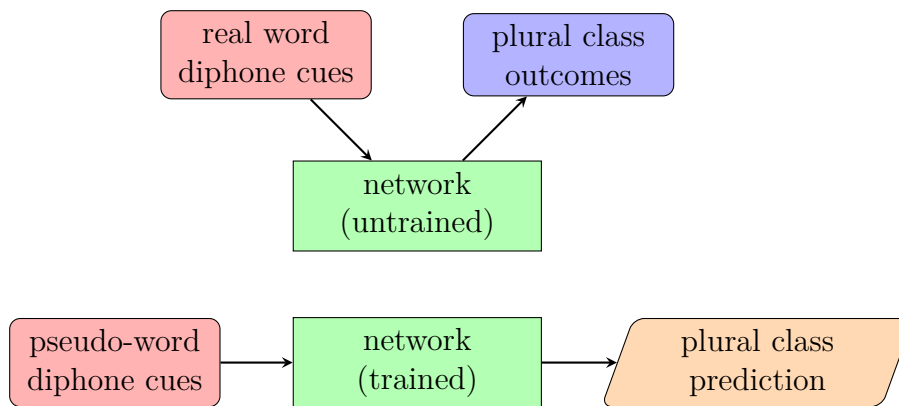


Figure 3: Training and test set up for **simple network design**. Top: training design; bottom: test design.

We tested to what extent the prediction of plural classes for pseudo-words varied depending on whether word form cues were based only on singular

word forms, on plural word forms, or on both, singular and plural word forms. A total of 3190 singular forms and 3190 plural forms were used. When cues were based on singulars and plurals, they were used as independent entries to maintain strong cue competition. Since Nieder, Tomaschek, et al. (2021) report a better classification of Maltese plural classes when word forms were transformed into diphones as cues in their NDL models, we used diphones in all our models.

Training of the model results in a network that represents the degree of connection strength between diphone cues and outcomes. In the second step, we tested what kind of plural classes the trained network predicted, when presented with the diphone cues of the pseudo-words used in Nieder, van de Vijver, et al.’s (2021a) wug test. This is accomplished by extracting the connection weights between each pseudo-word’s diphone cues and all plural class outcomes. By summing the weights between each cue set and each outcome, we obtain the pseudo-word’s activation. This measure represents the amount of support from a cue set to each outcome. The outcome with the highest activation – i.e. the plural class with the strongest support – is regarded as the winner of the classification process. Figure 3 illustrates these two steps for our modelling experiments.

We also tested a second modelling approach. This approach is based on the idea that speakers in a wug test do not directly ‘generate’ the inflected pseudo-word. Instead, the (orthographic/acoustic) cues of the pseudo-word activate a plural form of a real word. By analogy to the real plural word, the pseudo-word is inflected – which we classify into a plural class.

We tested this hypothetical process in a *double network* design, in which we trained two networks (network A and network B) in a first step and then made predictions about the possible plural class of the pseudo-word in the second step. As illustrated in Figure 4 (top left) network A was trained to discriminate real Maltese plural words on the basis of diphone cues of real Maltese words. As in the preceding network design, we used either singular nouns, plural nouns, or singular and plural nouns as a basis for diphone cues. Since in its current form NDL is not capable to generate any words on its own, we directly jumped to the plural class classification of the plural word by training Network B to discriminate plural classes on the basis of real Maltese plural words (Figure 4, top right).

To obtain predictions about the plural class for each pseudo-word (see Figure 4, bottom), we first used the diphone cues of the pseudo-word to obtain a prediction about a real plural word from network A. The diphone cues of the predicted plural word were in turn used to make obtain prediction

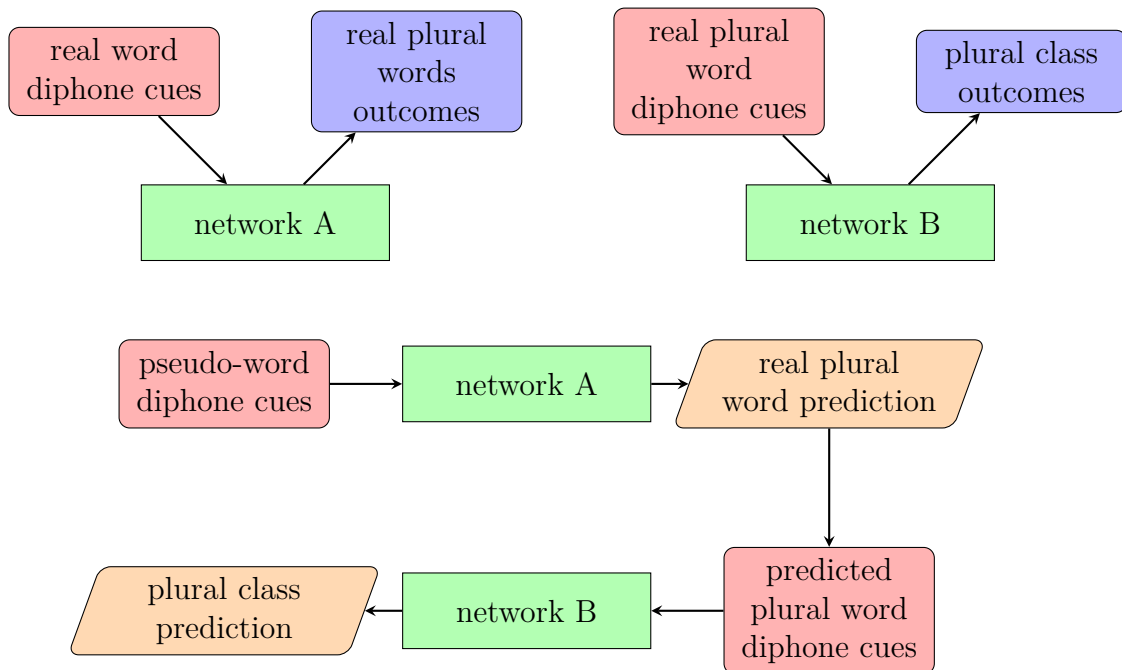


Figure 4: Training and test set up for **double network design**. Top: training design of the two networks A and B; mid: test design using both networks.

about the plural class from network B¹. The following list illustrates all tested set ups:

1. simple network: singular cues \rightarrow plural class
2. simple network: plural cues \rightarrow plural class
3. simple network: singular and plural cues \rightarrow plural class
4. double network: singular cues \rightarrow plural forms \rightarrow plural class

¹We also tested a ‘triple network design’ with a third network that was trained to discriminate singular word forms on the basis of diphones from singular nouns. In the test, the diphone cues of the pseudo-word are used to make a prediction about a real singular noun whose cues then are used to make a prediction about a real plural noun, which in turn predicts the plural class. This network design yielded worse results than the simple network and the double network design which is why we will not discuss it in detail in this study.

5. double network: plural cues \rightarrow plural forms \rightarrow plural class
6. double network: singular and plural cues \rightarrow plural forms \rightarrow plural class

Taking into account the different word forms used to create diphone cues (singulars, plurals, singular and plurals) and the two network designs, we ran six NDL models. Table 2 displays the cue-to-outcome structures for all six NDL models for the noun *kelb* ‘dog’ as an example. The Maltese word form *kelb* has a Semitic origin and shows the broken plural form *klieb* with its CV-pattern CCVVC. The diphone cues coded as quasi-phones are given in the fourth column of the table. Following the NDL notation, the hash mark # illustrates the beginning and end of a word form. The rightmost column displays the outcomes of the different models, whereas for the singular-plural models the outcome is based on a combination of the singular and plural cues (illustrated by combining both with curly brackets).

A note on the cue structure in our computational experiments: It is commonly assumed that morphological knowledge is knowledge of form-to-form mappings, with only a very rudimentary role of other linguistic modules, such as semantics. We follow this approach in the present paper. Obviously this is a very strong simplification of the speech production process. Speakers use language to make their intentions clear, i.e. by discriminating semantic contrasts, not to show off what versatile form-to-form mappers they are – exactly what they are being evaluated on in wug tests. Nevertheless we had two reasons not to include sophisticated semantics in our modelling efforts. First, in order to be able to compare rule-based and analogy-based approaches to the characterization of morphological knowledge, in which semantics also leads an atrophied life, it is necessary to ignore semantics for the time being. Second, it is far from trivial to obtain the meaning of pseudo-words, as has recently been demonstrated by Chuang et al. (2020). Since Nieder, van de Vijver, et al. (2021a)’s experiment consisted of relatively few pseudo-words (in comparison to the size of a corpus), we refrained from trying to include it in our current modelling study.

All material and models presented here can be found in our Supplementary Material at <https://osf.io/5es2q/>.

3.3 Comparing model predictions with wug test results

The aim of the present study was to investigate the knowledge Maltese native speakers demonstrated in a wug test on the basis of a discriminative learning approach. To do so, we test to what degree the use of plural classes of inflected pseudo-words in the wug test can be predicted by NDL. A direct

design	cues	word forms	diphone cues	outcomes
simple network	singular	<i>kelb</i> ‘dog’	#k, kE, El, lb, b#	broken CCVVC
	plural	<i>klieb</i> ‘dogs’	#k, kl, l8, 8b, b#	broken CCVVC
	singular-plural	<i>kelb</i> ‘dog’ <i>klieb</i> ‘dogs’	#k, kE, El, lb, b# #k, kl, l8, 8b, b#	} broken CCVVC
double network	singular	A. <i>kelb</i> ‘dog’ B. <i>klieb</i> ‘dogs’	#k, kE, El, lb, b# #k, kl, l8, 8b, b#	klieb broken CCVVC
	plural	A. <i>klieb</i> ‘dogs’ B. <i>klieb</i> ‘dogs’	#k, kl, l8, 8b, b# #k, kl, l8, 8b, b#	klieb broken CCVVC
	singular-plural	A. <i>kelb</i> ‘dog’ A. <i>klieb</i> ‘dogs’ B. <i>klieb</i> ‘dogs’	#k, kE, El, lb, b# #k, kl, l8, 8b, b# #k, kl, l8, 8b, b#	} klieb broken CCVVC

Table 2: Example of the cue-to-outcome structure for the noun *kelb* ‘dog’ in all six models during training. The cues are given in the quasi-phonetic transcription. The weights of these cues to the outcomes that were established in this way were used to assess the outcomes –plural classes– for the experimental stimuli.

comparison between the NDL predictions and the participants’ inflections is unfortunately not possible, the reason being that Nieder, van de Vijver, et al.’s (2021a) participants strongly vary in how exactly they inflect a pseudo-word. Concretely, one and the same pseudo-word may be inflected in different ways by different speakers (Heitmeier & Frank, 2021; Nieder, van de Vijver, et al., 2021a; van de Vijver & Baer-Henney, 2014). This suggests that – in contrast to the predictions of an abstraction approach – there is no truly ‘correct’ answer to how a pseudo-word has to be inflected. Since NDL will always make the same prediction for each pseudo-word, this means that NDL’s classification accuracy of plural classes for individual words will strongly vary between speakers. For example, the same item could obtain a *-i* suffix from one speaker and a *-ijiet* suffix from another one. One example

given by Nieder, van de Vijver, et al. (2021a) is the pseudo-word *follu* with its possible plurals *folol*, *folli* or *follijiet* provided by different participants. Instead of inspecting classification accuracies – which are typically considered an important criterion for a model (Albright & Hayes, 2003) –, we opted for an approach in which NDL can be regarded as a predictive mechanism for an entire community (Ernestus & Baayen, 2003; Hayes & Londe, 2006). From this perspective, what counts is not the predicted plural class for individual words, but the proportions of predicted plural classes in the entire data set of pseudo-words. Clearly, the proportions of plural classes will reflect that of the training corpus. However, by means of our network designs and different cue-to-outcome structures we changed the structure of inputs to the model and how the input is learned (see Section 3.2). Accordingly, we want to inspect which method best approximates the proportions of plural classes. This is accomplished by correlating the proportions of plural classes obtained by NDL and the proportions present in the corpus data as well as in Nieder, van de Vijver, et al.’s (2021a) wug experiment (see also Linke & Ramscar, 2020, for proportions-based approaches to investigating variation in languages).

4 Results

4.1 Correlations among speakers

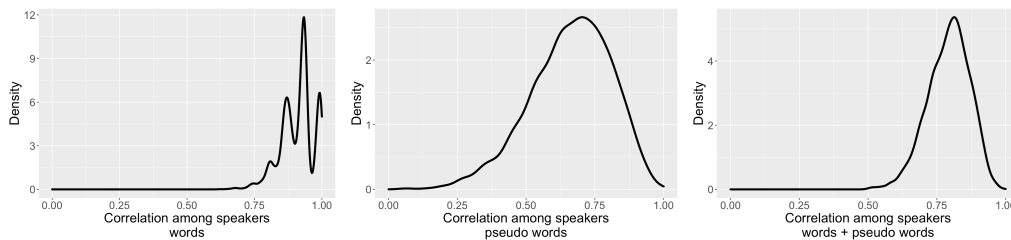


Figure 5: Proportions of correlations between plural class proportions for speakers in each condition: words, pseudo-words and words+pseudo-words

Before we present how the model performs in comparison to the corpus and the speakers, the question arises to what degree speakers agree in their use of plural classes in both the real words and the pseudo-words. To answer this question, we obtained a Spearman’s rank correlation between the proportions of plural classes for all pair-wise speaker combinations. Figure 5 illustrates the resulting distributions of correlations.

The left density plot illustrates that speakers agreed relatively consistently on the inflection of real words, i.e. words that they know and have potentially used before. This is indicated by a correlations that range mostly between $\rho = 0.75$ and $\rho = 1$, with four peaks at around 0.81, 0.88, 0.95 and 0.99. (These peaks correspond to the structure of the data: Nieder, van de Vijver, et al. (2021a) report that they included frequent sound and broken plurals and infrequent sound and broken plurals in their experiment, leaving us with four different conditions).

The center density plot illustrates the proportions of correlations between speakers for pseudo-words. We observe that the distribution of correlations ranges between $\rho = 0.2$ and $\rho = 1$, with most of the density mass located between $\rho = 0.5$ and $\rho = 1$ and a peak around 0.7. This indicates that while there was agreement about the plural classes between some participants, there were also pairs of participants who performed systematically differently in the wug test.

The right plot illustrates the correlation proportions when pseudo-words and real words were tested together, mirroring the effects of the two previous analyses.

In conclusion, we find that while speakers agree about the usage of plural classes for real words, they tend to disagree for non-words.

4.2 Correlations of proportions between NDL, corpus and wug experiment

The dotplots in Figure 6 illustrate the proportions (x-axis) of plural classes (y-axis) in the corpus (blue circles), in the participant responses in the wug test (pink triangles) and in the predictions obtained from NDL (grey squares). The left column illustrates the simple network design, the right column the double network design (the proportions for the corpus and the experiment are the same in both columns). The dotplots illustrate that overall there seems to be a strong correlation between the corpus, the experiment and the plural classes predicted by the two network designs. To evaluate this finding in more detail, we ran Spearman’s rank correlation tests to assess the relationship between the proportions in the corpus, the proportions in the experimental results and the proportions in the modelling approaches.

Unsurprisingly, we found a significantly positive correlation between the participants’ results and the proportions of the plural classes in the corpus ($\rho = .8603$, $p < 0.01$). As this correlation is the same in all three models, since all models use the same experimental and corpus data set, we will not report this correlation again for every NDL model.

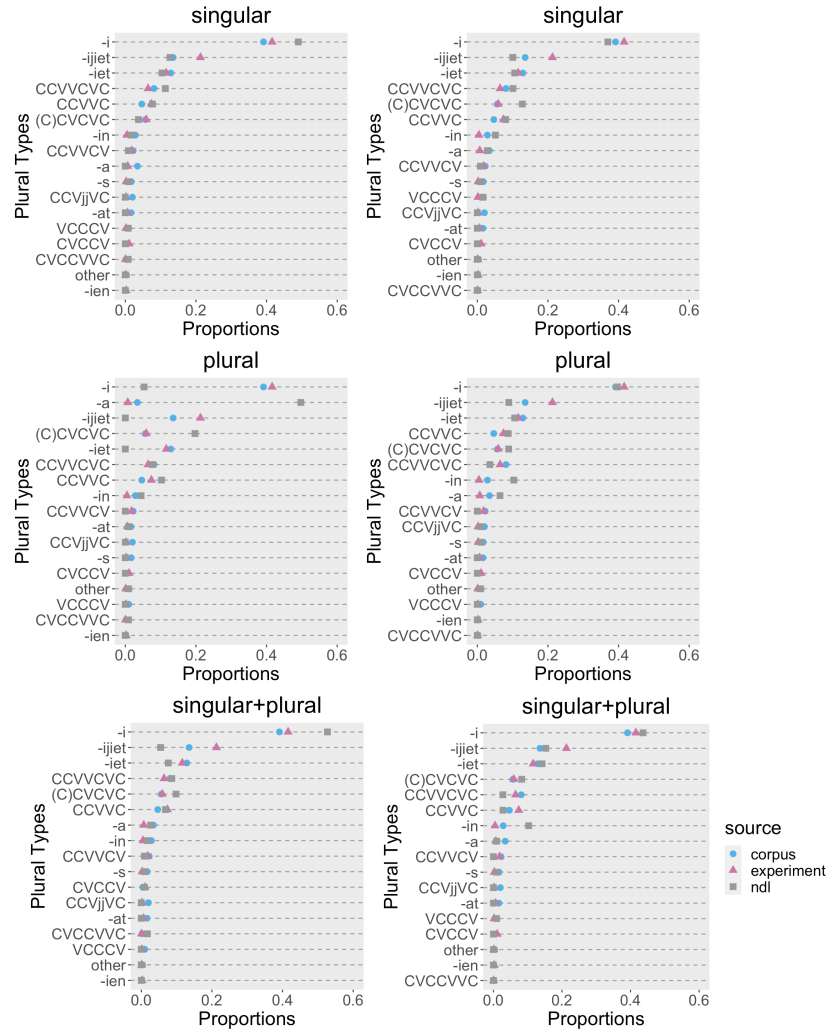


Figure 6: Proportion of plural classes used in the production experiment vs. plural classes used by the NDL model vs. plural classes in the corpus. **Left column:** simple network design. **Right column:** double network design.

Table 3 illustrates the correlations between the corpus, the results in the wug experiment and the proportions of predicted plural classes by NDL, depending on the cue structure provided to the network. Significant results are highlighted in bold font (for sake of clarity, we repeat the correlations here in the text again).

We first turn our attention to the findings depending on the simple network structure (top row). When using singulars as cues, we found a high positive correlation between all modelling results and the corpus data ($\rho = .820$, $p < 0.01$). This indicates that the model is capable to reflect the pro-

network	cues	corpus - NDL	experiment - NDL
Simple	singular	.820	.728
	plural	.353	.227
	singular-plural	.860	.759
Double	singular	.904	.768
	plural	.846	.650
	singular-plural	.866	.714

Table 3: Spearman’s rank correlations for all tested NDL models (displayed in Figures 4-9); significant correlations are highlighted in bold font.

portions of Maltese plural classes as attested in the corpus data. The high correlation between the participants’ results and the NDL data ($\rho = .728$, $p < 0.01$) illustrates that the simple model with singular cues is capable to predict relatively well the behavior of participants in a wug test.

When plurals were used as cues, we did not find any significant correlations between predicted proportions of plural classes and those in the corpus as well as in the experiment ($\rho = .353$, $p = 0.164$ / $\rho = .227$, $p = 0.3818$). The low, non-significant correlations indicate that plural forms are not sufficiently informative as to accurately predict the plural classes.

Instead, as displayed in Figure 3 (mid left), we find that NDL overgeneralizes the sound plural suffix *-a* and undergeneralizes the two most frequent suffixes *-i* and *-ijiet*.

When singulars and plurals are used as cues in the simple network design, we find that the correlations in both cases, the corpus and the experiment, are higher than when using singulars only ($\rho = 0.860$, $p < 0.01$ / $\rho = 0.759$, $p < 0.01$). This may come as a surprise, as the use of plural nouns as cues has not yielded significant results. However, recall that NDL takes into account cue-competition. This means that when plurals are used with singulars, together they very likely increase cue competition, making cues more informative about the plural classes. As a result, higher correlations are observed.

Next, we turn our attention to the correlations between the corpus and the experiment, and the proportions of plural classes predicted by the double network design. Here we find that using singulars as cues yields the highest correlations ($\rho = .904$, $p < 0.01$ / $\rho = .768$, $p < 0.01$) among the different cue structures provided to the double network design. Using plurals as cue yields

significant results. Nevertheless, plural cues result in lowest correlations ($\rho = .846$, $p < 0.01$ / significant, $\rho = .650$, $p < 0.05$), with singular and plurals as cues yielding mid effects ($\rho = .866$, $p < 0.01$ / $\rho = .714$, $p < 0.01$).

5 Discussion and conclusion

5.1 Summary

As the complexities of Maltese nouns puts them beyond the purview of abstraction-based approaches, and analogy-based approaches make unwarranted assumptions about the assessment of similarity and learning, we decided to make demands on the Discriminative Lexicon by using it to model the results of a Maltese wug test.

Our aim was to shed light on the morphological knowledge speakers use to inflect pseudo-words in wug tests. To this end, we used NDL (Baayen et al., 2011; Shaoul et al., 2013), a two-layer network rooted in Discriminative Learning (Ramscar et al., 2010) to model the proportions of plural classes of inflected pseudo-words in Nieder, van de Vijver, et al. (2021a)’s wug test with 80 adult native speakers of Maltese.

We ran NDL models with different cue-to-outcome structures in which the plural class was predicted either on the basis of a singular form, a plural form or both. Moreover, we tested two network designs to make predictions for the wug test. In the first, cues of real word forms predicted plural classes directly; in the second, called ‘double network’, cues of real word forms first predicted real plural words, whose cues in turn predicted the plural classes (see Figures 3 and 4 again for a detailed description of the designs).

Once we obtained the plural class predictions for each pseudo-word, we compared the proportions of the plural classes predicted by NDL to the proportions of the plural classes in the wug test, and to the proportions of the plural classes in the Maltese corpus. Comparisons were performed by means of a Spearman’s rank correlation test.

As illustrated in Table 3, we find that the highest correlations between plural class proportions were obtained when both, singulars and plurals were used as cues in the training in the simple network design. The lowest – and thus not significant – correlations were obtained when the model was provided with plural cues only. The question thus arises why this is the case.

One reason for why using plural cues yielded no significant correlations may be found in similarities between the word forms of real plurals and pseudo singulars in Nieder, van de Vijver, et al.’s (2021a) study. In the training phase, NDL had only access to plural cues and learned to relate

them to plural class outcomes. Some of the words with an offset *-a* are broken plurals that simply end in the vowel. However, many words actually belong to a sound plural class in which the plurals end in *-a*. Since *-a* is the Maltese feminine marker for nouns, Nieder, van de Vijver, et al. (2021a) constructed their pseudo words such that many of the pseudo singulars end in an *-a* (Borg & Azzopardi-Alexander, 1997). When faced with the pseudo-words, but trained on real plurals, this learned relation most likely forced NDL to predict the sound plural *-a* for a great amount of pseudo singulars that end in *a*. NDL overgeneralized the sound plural suffix *-a* and undergeneralized the sound plural suffix *-i* (see Figure 6, bottom left) in the pseudo singulars used by Nieder, van de Vijver, et al.’s (2021a) study. As a result, we observe a low correlation.

Turning our attention to the double network design, we find that using singular cues yielded the highest correlation while adding plural word form to the training reduced the correlation. In addition, the proportions of plural classes obtained with the double network had a higher correlation with the proportions of plural classes in both the corpus and the experiment. These findings have implications for our understanding of wug tests that we will discuss below.

5.2 Implications

The high correlations between the proportions of predicted plural classes and the proportions in the experiment (Table 3) demonstrates that our model is capable of modeling the behavior of speakers in Nieder, van de Vijver, et al.’s (2021a) wug experiment. This, obviously, has implications for theories of morphology and the mental lexicon.

Since production of novel inflected forms can be modeled without recourse to morphemes, an implication for morphological theory is that morphology arises from the mapping of form onto function (as in our model) or meaning (as in Baayen et al. (2019)). Recourse to abstract categories, such as morphemes, is unnecessary (see also Ambridge, 2020). This finding is in turn in line with other work on Maltese nouns (Nieder, Tomaschek, et al., 2021; Nieder, van de Vijver, et al., 2021a, 2021b), and is explained by the theory of Word and Paradigm (Blevins, 2016).

Our current approach, then, is in agreement with much recent work arguing that morphemes, or other abstract categories, are maybe helpful in typological or diachronic analyses of languages, but are not part of the knowledge that native speakers have of their language (Ambridge, 2020). The rules that manipulate prosodic structure as morphemes and which have been proposed to account for Maltese broken plurals (Schembri, 2012) succinctly describe

groups of broken plurals. But our modelling shows that it is possible to classify Maltese plurals into groups without rules or prosodic structure, or morphemes. As it is not clear how the rules would apply to novel words, our modelling is a better fit to the content of the mental lexicon of Maltese native speakers.

Our model is part of the theory of the Discriminative Lexicon in which the associations between word forms and their grammatical functions are learned in an error-driven way. The modelling in the Discriminative Lexicon theory is based on error-driven learning which is well-supported cognitively for language learning and processing (Nixon, 2020; Nixon & Tomaschek, 2021; Olejarczuk et al., 2018; Ramskar, Dye, & Klein, 2013; Ramskar & Yarlett, 2007) and offers an alternative to analogical modelling. This is necessary, amongst other reasons, because in analogical modelling it is unclear how similarity is learned. In these analogical models, the start of the similarity comparison is pre-specified by the researcher, but it is unclear how a human learner can decide what is similar to what, why and where to start (Gahl & Strand, 2016).

Our results raise the question: Can we, on the basis of the present models and results, infer any kind of implications for cognitive processes that may take place during pseudo-word inflection? While we succeeded in modelling the plural class proportions, the modelling designs we applied do not very likely reflect any type of cognitive process happening in native speakers of Maltese (or in a broad sense of any language). In the simple network design, we mapped phonological word forms onto plural classes that are the result of grammatical analyses and of which speakers of Maltese are most likely unaware (or if they are, they are the result of long school education). This problem applies more so to the double network design in which the classification took a detour over a real plural word. Speakers do not actively select a plural word on the basis of the pseudo-word and assign a plural class to it.

Another perspective missing in the present approach is the role of semantics. Semantic information is a prerequisite for inflectional classes in many languages (Haspelmath, 2013) and recent modelling studies demonstrated that it improves prediction accuracies when taken into account also for languages such as German or English (e.g. Baayen et al., 2019). Ramskar (2019) even argues that semantics – even in the form of inflectional functions which establish semantic relations in a given context – are the cues which discriminate word forms. This perspective has recently been shown to result in better predictions of fine phonetic detail than when functions are the outcome of the discriminative process (Tomaschek & Ramskar, *submitted*). By contrast, the present classification was based purely on the phonological properties of the words and ignored potential interference from sentential context.

However, each of these arguments could be used against any kind of computational modelling approach that was applied in the last decade to investigate cognitive processes and their effects on perception and production. As all computational models, the present approach was a simplification of speech production which in reality is a highly complex neural process that results from interactions of different layers of abstraction (Hickok, 2014). Accepting this reduction of complexity as a valid step, we regard our approach to be successful in demonstrating that inflecting pseudo-words does not necessarily need to be based on rules. Instead, our results indicate that inflection of pseudo-words in wug tests – and therefore very likely the inflection of new words encountered in every day language – is based on analogical processes. Speakers are capable to apply these analogical processes on the basis of their (learned) knowledge of the relations between the phonological properties of inflected words and the grammatical functions they want to express (such as plural). When presented with pseudo-words, together with the ‘plural’ cue that is implicitly present in the experiment, the words’ phonological cues start a discriminative, predictive process during which either the abstraction in terms of inflectional process (represented by our plural class) is started. Or, as indicated by the double network design, the cues of the pseudo-word first activate real inflected words that serve as a template for the inflection of the pseudo-word. At no point in time there is a rule. Instead inflection of pseudo-words is based on cue-to-outcome structures that are learned. The question thus arises what the responsible mechanism is that allows speakers to learn the association between phonological form and inflectional type? From our theoretical perspective, the knowledge that speakers use to inflect pseudo-words emerges through discriminative learning.

In conclusion, the responses in wug tests do not need to be explained by assuming the application of morphological and phonological rules to morphemes. Instead, speakers learn how sounds discriminate among meanings and functions of words, and apply this knowledge to pseudo-words.

Acknowledgement & Funding

This research was funded by the Deutsche Forschungsgemeinschaft (Research Unit FOR2373 ‘Spoken Morphology’, Projects VI 223/3-2 and BA 3080/3-2), which we gratefully acknowledge.

We thank the audience of the Colloquium of the research unit Spoken Morphology as well as the audience of the Words in the World International

Conference 2021 for their feedback and comments.

Data Availability Statement

The data that support the findings of this study are openly available at <https://osf.io/5es2q/>.

References

- Albright, A., Andrade, A., & Hayes, B. (2001). Segmental environments of Spanish diphthongization. *UCLA Working Papers in Linguistics*, 7, 117–151.
- Albright, A. (2002a). A restricted model of UR discovery: Evidence from Lakhotá. *Ms, University of California at Santa Cruz*.
- Albright, A. (2002b). Islands of reliability for regular morphology: Evidence from Italian. *Language*, 78, 684–709.
- Albright, A., & Hayes, B. (2003). Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition*, 90(2), 119–161. [https://doi.org/10.1016/S0010-0277\(03\)00146-X](https://doi.org/10.1016/S0010-0277(03)00146-X)
- Ambridge, B. (2020). Against stored abstractions: A radical exemplar model of language acquisition. *First Language*, 40(5-6), 509–559. <https://doi.org/10.1177/0142723719869731>
Publisher: SAGE Publications Ltd
- Arnold, D., Tomaschek, F., Sering, K., Lopez, F., & Baayen, R. H. (2017). Words from spontaneous conversational speech can be recognized with human-like accuracy by an error-driven learning algorithm that discriminates between meanings straight from smart acoustic features, bypassing the phoneme as recognition unit. *PLOS ONE*, 12(4), e0174623. <https://doi.org/10.1371/journal.pone.0174623>
- Arppe, A., Hendrix, P., Milin, P., Baayen, R. H., Sering, T., & Shaoul, C. (2018). Ndl: Naive Discriminative Learning. Retrieved November 20, 2020, from <https://CRAN.R-project.org/package=ndl>
- Baayen, R. H., Milin, P., Durdevic, D. F., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological review*, 118(3), 438–481.
- Baayen, R. H. (2011). Corpus linguistics and naive discriminative learning. *Brazilian Journal of Applied Linguistics*, 11, 295–328.

- Baayen, R. H., Chuang, Y.-Y., Shafaei-Bajestan, E., & Blevins, J. P. (2019). The discriminative lexicon: A unified computational model for the lexicon and lexical processing in comprehension and production grounded not in (de) composition but in linear discriminative learning. *Complexity*, 2019.
- Baayen, R. H., Hendrix, P., & Ramscar, M. (2013). Sidestepping the combinatorial explosion: An explanation of n-gram frequency effects based on naive discriminative learning. *Language and Speech*, 56(3), 329–347.
- Baayen, R. H., & Smolka, E. (2020). Modeling morphological priming in German with naive discriminative learning. *Frontiers in Communication*, 5, 17.
- Berko, J. (1958). The Child’s Learning of English Morphology. *Word*, 14, 150–177.
- Blevins, J. P. (2016). *Word and paradigm morphology*. Oxford University Press.
- Borg, A., & Azzopardi-Alexander, M. (1997). *Maltese*. Routledge.
- Bybee, J. (2010). *Language, usage and cognition*. Cambridge University Press.
- Camilleri, J. J. (2013). *A Computational Grammar and Lexicon for Maltese*. [Master thesis, University of Gothenburg].
- Chomsky, N. (1957). *Syntactic Structures*. Mouton.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT press.
- Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. Harper; Row.
- Chuang, Y.-Y., Lõo, K., Blevins, J. P., & Baayen, R. (2019). Estonian case inflection made simple. A case study in Word and Paradigm morphology with Linear Discriminative Learning.
- Chuang, Y.-Y., Vollmer, M. L., Shafaei-Bajestan, E., Gahl, S., Hendrix, P., & Baayen, R. H. (2020). The processing of pseudoword form and meaning in production and comprehension: A computational modeling approach using linear discriminative learning. *Behavior Research Methods*. <https://doi.org/10.3758/s13428-020-01356-w>
- Daelemans, W., & Van den Bosch, A. (2005). *Memory-based language processing*. Cambridge University Press.
- Danks, D. (2003). Equilibria of the Rescorla–Wagner model. *Journal of Mathematical Psychology*, 47, 109–121.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological review*, 93(3), 283–321.
- Divjak, D., Milin, P., Ez-zizi, A., Józefowski, J., & Adam, C. (2020). What is learned from exposure: an error-driven approach to productivity in

- language. *Language, Cognition and Neuroscience*, 36(1), 1–24. <https://doi.org/10.1080/23273798.2020.1815813>
- Ellis, N. C. (2006a). Language Acquisition as Rational Contingency Learning. *Applied Linguistics*, 27(1), 1–24. <https://doi.org/10.1093/applin/ami038>
- Ellis, N. C. (2006b). Language acquisition as rational contingency learning. *Applied linguistics*, 27(1), 1–24.
- Ernestus, M., & Baayen, R. H. (2003). Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language*, 79, 5–38.
- Fromkin, V. A. (1971). The Non-Anomalous Nature of Anomalous Utterances. *Language*, 47(1), 27–52. <https://doi.org/10.2307/412187>
 Publisher: Linguistic Society of America
- Fromkin, V. A. (1988). Grammatical aspects of speech errors. *Linguistics: The Cambridge Survey*, 2, 117–138.
- Gahl, S., & Strand, J. F. (2016). Many neighborhoods: Phonological and perceptual neighborhood density in lexical production and perception. *Journal of Memory and Language*, 89, 162–178.
- Goldsmith, J. (1979). *Autosegmental phonology*. [Doctoral dissertation, MIT].
- Haspelmath, M. (2013). 34 occurrence of nominal plurality. *The World Atlas of Language Structures Online*.
- Hayes, B., & Londe, Z. C. (2006). Stochastic phonological knowledge: The case of Hungarian vowel harmony. *Phonology*, 23(1), 59–104.
- Heitmeier, M., & Frank, S. (2021). Exploring the influence of semantics on the German plural system: a wug study. *Proceedings of the Annual Meeting of the Cognitive Science Society*.
- Hickok, G. (2014). Towards an integrated psycholinguistic, neurolinguistic, sensorimotor framework for speech production. *Language, Cognition and Neuroscience*, 29(1), 52–59. <https://doi.org/10.1080/01690965.2013.852907>
- Indefrey, P., & Levelt, W. J. (2000). The neural correlates of language production. *The new cognitive neurosciences; 2nd ed.*, 845–865.
- Keuleers, E., Sandra, D., Daelemans, W., Gillis, S., Durieux, G., & Martens, E. (2007). Dutch plural inflection: The exception that proves the analogy. *Cognitive Psychology*, 54(4), 283–318.
- Kunter, G. (2017). Coquery: A free corpus tool. version 0.10.0.. <https://www.coquery.org>
- Leben, W. R. (1973). *Suprasegmental phonology*. [Doctoral dissertation, Massachusetts Institute of Technology].
- Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *The Behavioral and brain sciences*, 22(1).

- Linke, M., & Ramscar, M. (2020). How the Probabilistic Structure of Grammatical Context Shapes Speech. *Entropy*, *22*(1), 90.
- Marantz, A. (1982). Re reduplication. *Linguistic Inquiry*, *13*(3), 435–482. <http://www.jstor.org/stable/4178287>
- Marcus, G. F., Brinkman, U., Clahsen, H., Wiese, R., & Pinker, S. (1995). German inflection: The exception that proves the rule. *Cognitive Psychology*, *29*, 189–256.
- Marcus, G. F., Brinkmann, U., Clahsen, H., Wiese, R., & Pinker, S. (1995). German inflection: The exception that proves the rule. *Cognitive psychology*, *29*(3), 189–256. <https://doi.org/https://doi.org/10.1006/cogp.1995.1015>
- Martinet, A. (1965). *La Linguistique Synchronique: \textbackslash'Etudes et Recherches*. Presses Universitaires de France.
- McCarthy, J. J., & Prince, A. S. (1990). Foot and word in prosodic morphology: The Arabic broken plural. *Natural Language and Linguistic Theory*, *8*(2), 209–283.
- Milin, P., Feldman, L. B., Ramscar, M., Hendrix, P., & Baayen, R. H. (2017). Discrimination in lexical decision. *PLOS ONE*.
- Nieder, J., Tomaschek, F., Cohrs, E., & van de Vijver, R. (2021). Modelling Maltese noun plural classes without morphemes. *Language, Cognition and Neuroscience*, 1–22. <https://doi.org/10.1080/23273798.2021.1977835>
- Nieder, J., van de Vijver, R., & Mitterer, H. (2021a). Knowledge of Maltese singular-plural mappings. Analogy explains it best. *Morphology*, *31*, 147–170. <https://doi.org/https://doi.org/10.1007/s11525-020-09353-7>
- Nieder, J., van de Vijver, R., & Mitterer, H. (2021b). Priming Maltese Plurals: Representation of sound and broken plurals in the mental lexicon. *Mental Lexicon*, *16*(1), 69–97. <https://doi.org/https://doi.org/10.1075/ml.20008.nie>
- Nixon, J. S. (2020). Of mice and men: Speech sound acquisition as discriminative learning from prediction error, not just statistical tracking. *Cognition*, *197*, 104081. <https://doi.org/https://doi.org/10.1016/j.cognition.2019.104081>
- Nixon, J. S., & Tomaschek, F. (2021). Prediction and error in early infant speech learning: A speech acquisition model. *Cognition*, *212*, 104697.
- Olejarczuk, P., Kapatsinski, V., & Baayen, R. H. (2018). Distributional learning is error-driven: The role of surprise in the acquisition of phonetic categories. *Linguistics Vanguard*, *4*(s2).

- Pham, H., & Baayen, R. H. (2015). Vietnamese compounds show an anti-frequency effect in visual lexical decision. *Language, Cognition and Neuroscience*, 30(9), 1077–1095.
- Pinker, S. (1989). Learnability and cognition. The acquisition of argument structure.
- Pinker, S., & Prince, A. (1991). Regular and irregular morphology and the psychological status of rules of grammar. *Proceedings of the 1991 meeting of the Berkeley Linguistics Society*.
- Pinker, S., & Prince, A. (1988). On language and connectionism: Analysis of parallel distributed processing model of language acquisition. *Cognition*, 28(1-2), 73–193. [https://doi.org/https://doi.org/10.1016/0010-0277\(88\)90032-7](https://doi.org/https://doi.org/10.1016/0010-0277(88)90032-7)
- Prince, A., & Smolensky, P. (2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Blackwell.
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. <https://www.R-project.org/>
- Ramscar, M., Dye, M., & Klein, J. (2013). Children Value Informativity Over Logic in Word Learning -. *Psychological Science*, 24(6), 1017–1023. <https://doi.org/https://doi.org/10.1177/0956797612460691>
- 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. (2007). Linguistic Self-Correction in the Absence of Feedback: A New Approach to the Logical Problem of Language Acquisition. *Cognitive Science*, 31(6), 927–960.
- 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.
- Ramscar, M. (2019). *Source codes in human communication* (preprint). PsyArXiv. <https://doi.org/10.31234/osf.io/e3hps>
- Rescorla, R. (1988). Pavlovian conditioning: It’s not what you think it is. *American psychologist*, 43(3), 151.
- Rescorla, R., & Wagner, A. (1972a). A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–69). Appleton Century Crofts, New York.
- Rescorla, R., & Wagner, A. (1972b). 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.

- Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, 64(3), 249–284.
- Roelofs, A., & Ferreira, V. S. (2019). The architecture of speaking. In P. Hagoort (Ed.), *Human language: From genes and brains to behavior* (pp. 35–50). MIT Press Cambridge.
- Rumelhart, D. E., & McClelland, J. L. (1986). On learning the past tenses of English verbs. In G. T. M. Altmann (Ed.), *Psycholinguistics*.
- Schembri, T. (2012). *The Broken Plural in Maltese: A Description*. Univ.-Verl. Brockmeyer.
- Schmitz, D., Plag, I., Baer-Henney, D., & Stein, S. D. (2021). Durational differences of word-final/s/emerge from the lexicon: Modelling morpho-phonetic effects in pseudowords with linear discriminative learning. *Frontiers in Psychology*, 12, 2983.
- Shafaei-Bajestan, E., Moradipour-Tari, M., Uhrig, P., & Baayen, R. H. (2020). *LDL-AURIS: Error-driven Learning in Modeling Spoken Word Recognition* (tech. rep.). PsyArXiv. <https://doi.org/10.31234/osf.io/v6cu4>
- Shaoul, C., Arppe, A., Hendrix, P., Milin, P., & Baayen, a. R. H. (2013). *Ndl: Naive Discriminative Learning*. Retrieved January 22, 2014, from <http://cran.r-project.org/web/packages/ndl/index.html>
- Skousen, R. (1989). *Analogical Modeling of Language*. Kluwer.
- Skousen, R. (1992). *Analogy and Structure*. Springer Netherlands.
- Tomaschek, F., Plag, I., Ernestus, M., & Baayen, R. H. (2019). Phonetic effects of morphology and context: Modeling the duration of word-final S in English with naïve discriminative learning. *Journal of Linguistics*, 57(1), 123–161.
- Tomaschek, F., & Ramscar, M. (submitted). Understanding the phonetic characteristics of speech under uncertainty – Implications of the representation of linguistic knowledge in learning and processing.
- Tucker, B. V., & Tomaschek, F. (to appear). Speech Production: Where does morphology fit? In D. Crepaldi (Ed.), *Current Issues in the Psychology of Language*. Routledge.
- van de Vijver, R., & Baer-Henney, D. (2014). Developing biases. *Frontiers in Psychology - Section Language Sciences*, 5, Article 634. <https://doi.org/https://doi.org/10.3389/fpsyg.2014.00634>
- Woo, N. (1969). Prosodic phonology. *Unpublished Doctoral dissertation, MIT*.

Quasi-phonetic transcription

phon	SAMPA	Example
p	p	<u>p</u> ajjiz
b	b	<u>b</u> ank
t	t	<u>t</u> abib
d	d	<u>d</u> enfil
k	k	<u>k</u> arrozza
g	g	<u>g</u> araxx
m	m	<u>m</u> aratona
n	n	<u>n</u> orma
l	l	<u>l</u> ibsa
r	r	<u>r</u> agel
f	f	<u>f</u> jura
v	v	<u>v</u> otant
s	s	lang <u>s</u> a
z	z	<u>z</u> ur <u>z</u> ieqa
S	S	<u>x</u> ena
j	j	vin <u>j</u> a
h	h	<u>h</u> obz
=	ts	<u>z</u> alzett
J	tS	<u>č</u> umnija
G	dZ	<u>g</u> urnata
i	i:	kn <u>i</u> sja
u	u:	mu <u>r</u>
)	E:	fe <u>h</u> em
X	OI	ero <u>j</u>
I	I	fi <u>t</u> t
E	E	fe <u>t</u> tet
O	O	so <u>d</u> da
U	U	ku <u>n</u> tratt
w	w	wi <u>z</u> għa
\$	O:	ro <u>g</u> ħda
?	?	q <u>a</u> lb
6	6	ra <u>s</u> s
9	6:	għa <u>n</u> i
8	I:	zi <u>e</u> mel
1	6U	aw <u>g</u> uri
2	6I	aj <u>r</u> uplan
3	EU	te <u>w</u> a
4	EI	rgħi <u>b</u> a
5	IU	ti <u>w</u> i
7	OU	go <u>w</u> l

Table 4: Rules for quasi-phonetic transcription, table taken from Nieder, Tomaschek, et al. (2021)

Mathematical model of error-driven, discriminative learning

The NDL model is formally a simple 2-layer network, in which a layer of cues is linked to a layer of outcomes. Using the Rescorla-Wagner learning equations (Rescorla, 1988; Rescorla & Wagner, 1972b), the network calculates association strength V_i^{t+1} between a cue C_i and an outcome O at time $t + 1$ as calculated in equation 4.

$$V_i^{t+1} = V_i^t + \Delta V_i^t \quad (4)$$

The change in association strength between the cues and outcomes is weighted by the equations in 5.

$$\Delta V_i^t = \begin{cases} 0 & \text{if } Absent(C_i, t) \\ \alpha_i \beta_i (\lambda - \sum_{Present(C_j, t)} V_j) & \text{if } Present(O, t) \\ \alpha_i \beta_i (0 - \sum_{Present(C_j, t)} V_j) & \text{if } Absent(O, t) \end{cases} \quad (5)$$

This model has been computationally implemented in the R environment (R Core Team, 2020) in the `ndl` package (Arppe et al., 2018). To avoid a-priori chunking of the words, we have used the words forms chunked into diphones as cues and plural classes as output. In the present study, we used the Danks Equilibrium Equations (Danks, 2003), which estimate the association weights for a stable state of the learning system (as implemented in the NDL (Arppe et al., 2018) in R (R Core Team, 2020)).