A Discriminative Lexicon Approach to Word Comprehension, Production and Processing: Maltese Plurals

Abstract

Comprehending and producing words is a natural process for human speakers. In linguistic theory, investigating this process formally and computationally is often done by focusing on forms only. By moving beyond the world of forms, we show in this study that the Discriminative Lexicon (DL) model operating with word comprehension as a mapping of form onto meaning and word production as a mapping of meaning onto form generates accurate predictions about what meanings listeners understand and what forms speakers produce. Furthermore, we show that measures derived from the computational model are predictive for human reaction times. Although mathematically very simple, the linear mappings between form and meaning posited by our model are powerful enough to capture the complexity and productivity of a Semitic language with a complex hybrid morphological system.

Keywords Discriminative Lexicon; Maltese Plurals; Word and Paradigm Morphology; Linear Discriminative Learning; Computational Modeling; Productivity; Primed Lexical Decision

1 Introduction

Most formal and computational accounts of word structure unfold almost exclusively in the world of forms: Forms are related to and compared with other forms. For instance, the prosodic theory of non-concatenative morphology laid out in McCarthy (1981) starts with underlying forms that are the starting point for a set of rules that derive words' surface forms. A Semitic verb form is conceived of as consisting of a root, filled with consonants which carry the meaning of the lexeme and its derivations, and a melody of vowels in which inflection is expressed. The Arabic form *kataba* 'he wrote' consists of the root $\sqrt{ktb}$, which expresses the lexeme 'to write' and the melody *aaa*, which expresses *third person singular past*. Both, consonantal root and vowel melody are mapped onto a skeleton CVCVCV from left to right resulting in the final word form *kataba*. Nouns, too, can have non-concatenative inflections; in Arabic or Maltese non-concatenative plurals are referred to as broken plurals. These plurals can be analyzed prosodically.

In subsequent work, the nature of the CV-skeleton changed, but not the fact that forms are mapped onto forms. McCarthy and Prince (1990, 1996) developed a theory, called Prosodic Morphology, in which the skeleton is replaced by prosodic categories using Arabic non-concatenative broken plurals as a testing ground. The singular *nafs* 'soul' has a corresponding broken plural *nufuus*. This plural can be characterized as an iamb: A light syllable followed by a heavy syllable. Formally McCarthy and Prince (1990) account for the iambic plural as a mapping of the phonological material from the leftmost superheavy syllable (*nafs*) of the singular onto an iamb and a concomitant change of vowel quality.

In a further attempt to reduce stipulations about the shape of non-concatenative morphology, Kastner (2019) proposed that the symbols for the verb and its inflectional features are first inserted into a syntactic tree. Subsequently, general principles of the Hebrew sound system account for the arrangement of the segmental material of the verbal root and its inflectional exponents. For instance, the root node $\sqrt{ktb}$ 'to write' and its voice specification {a,a} for past tense are inserted into the syntactic tree in a concatenative fashion as [Tense[Past [Voice [v $\sqrt{ktb}$]]]] (where small v is a functional head). This results in the input form *ktb,aa* for the phonological component. A hierarchy of constraints (Prince & Smolensky, 2004) then predicts the optimal output *katab*. Instead of deriving Hebrew verbal forms from consonantal roots, as argued by Kastner (2019) and McCarthy (1981), Ussishkin (2005) proposes that words are derived from other words, subject to a set of prosodic and morphological constraints.

Many computational models of morphology likewise do not predict words' forms from their meanings, but from other forms of these words. Some of these models set up a list of possible changes that have to be applied to move from one form to another, and then seek to predict which of the possible form changes is appropriate given selected properties of the base word. For instance, Ernestus and Baayen (2003) examined several quantitative models that all were given the task to predict whether or not the stem-final obstruent of a Dutch plural noun or verb form is voiced or voiceless. These models, which ranged from recursive partitioning trees and logistic regression models to Analogical Modeling (Skousen, 1989), Memory-Based learning (Daelemans & Van den Bosch, 2005) and Optimality Theory (Boersma & Hayes, 2001), all performed with roughly the same accuracy, suggesting that any reasonably decent statistical classifier, given access to the relevant features of the base word, can accomplish this classification task.[1] However, all these models are incomplete, in the sense that to create an actual plural form, the appropriate voicing feature has to be combined with further concatenation of the appropriate plural suffix.

For Semitic languages such as Arabic and Maltese, predicting the plural of a noun is set up as a classification problem by Dawdy-Hesterberg and Pierrehumbert (2014), focusing on Arabic, and by Nieder, Tomaschek, et al. (2021), focusing on Maltese. The former study used the Generalized Context Model (Nosofsky, 1986), the latter study applied Memory-Based learning (Daelemans et al., 2001), Naive Discriminative Learning (Baayen, 2011), as well as an Encoder-Decoder deep learning architecture (McCoy et al., 2020) to generate plurals from singulars. The deep learning model stands in the tradition of the past-tense model of Rumelhart and McClelland (1986), who derived English past-tense forms from their present-tense counterparts.

The only way in which semantics plays a role in these grammatical and computational models of inflection is through inflectional contrasts, such as singular versus plural, which are used to set up separate classes of forms (Albright & Hayes, 2003). However, it seems unlikely that native speakers produce plurals from singulars (or vice versa). In her classic study of the knowledge of children of English morphology, Berko (1958) notes that only 28% of the 4- and 5-year olds, and only 38% of the 5- and 7-year olds provided the plural *gutchess* for the given singular *gutch*. Van de Vijver and Baer-Henney (2014) also found that many children repeat a novel given singular as plural in a wug test in German. Zamuner et al. (2006) found that Dutch children are unable to form a novel singular from a given, novel plural, while they have no problems providing a singular from a known plural. How-

[1] Thus, the recursive partitioning algorithm of Belth et al. (2021) is also likely to perform well.

ever, they were reasonably successful in providing novel plurals from novel singulars. Klafehn (2013) reports that Japanese native speakers are unable to provide inflected words for provided novel words.

All these results suggest that producing a novel word form is not necessarily as straightforward as simply applying a rule to manipulate a form in an appropriate context. But it is not just results from wug tests that strongly indicate that plurals are not formed on the basis of their singulars. Bybee (1995) argues that in Hausa it is more insightful to characterize a plural in a product-oriented way: as a characterization of what makes a good plural in Hausa, rather than as an instruction how to turn a singular into a plural. This is because there is little predictability from singular to plural. A plural and a singular consist of overlapping phonological material, but the overlap can be inconsistent from singular-plural pair to singular-plural pair. Of course, there are situations in which form-to-form mappings are useful. For instance, instructions for how to create the forms of a paradigm from its principal parts can be quite helpful for second language learners, as a way to efficiently master paradigms, and for analysts, as a way of coming to grips with the systems of implications among word forms. But whether native speakers derive forms via other forms remains an open question (Blevins, 2016; Nieder, Tomaschek, et al., 2021).

In this study, we move beyond the world of forms, and model comprehension as a mapping of form to meaning and production as a mapping of meaning to form. We make use of a computational implementation of Word and Paradigm Morphology (Blevins, 2016; Matthews & Matthews, 1972), the 'discriminative lexicon' (DL) (Baayen et al., 2019), to model the noun system of Maltese, a Semitic language spoken in Europe. The DL model differs from most theories of morphology in that comprehension and production is achieved without requiring theoretical constructs such as stems, exponents, and inflectional classes. In general, the task of morphological theory is often conceptualized as providing a formal mechanism specifying what sound sequences are possible meaningful words. The DL model divides this task into two subtasks: first, to predict what possible forms are, given their meanings; and second, to predict what possible meanings are, given their forms.

In psychology, several computational models have been put forward that construct complex words starting from their meanings. The models by Levelt et al. (1999) and Dell (1986) are similar in design to realizational theories of morphology (see, e.g., Bonami & Stump, 2016; Stump, 2001). To our knowledge, these two psychological computational models have not been implemented for and applied to languages other than English, and it is therefore unclear whether the mechanisms of spreading activation and interactive activation, that they make use of, can be made to work for complex morphological systems such as the Maltese noun system.

Gaskell and Marslen-Wilson (1997) proposed a three-layer network model that maps speech input onto semantics, while explicitly shying away from making claims about representations that might develop in the hidden layer of their network. The triangle model of Harm and Seidenberg (2004) likewise addresses the relation between words' forms and their meanings, using a more complex multi-layer network. This model has been tested not only on English, but also on Serbo-Croatian (Mirković et al., 2005). Following their lead, the 'Discriminative Lexicon' model (Baayen et al., 2019) zooms in on the mappings from form to meaning in visual and auditory comprehension, and the mapping from meaning to form in production. As in the above connectionist models, both words' forms and their meanings are represented by high-dimensional numeric vectors. However, the DL model simplifies the connectionist multi-layer networks of Gaskell and Marslen-Wilson (1997) and Harm and Seidenberg (2004) by removing all hidden layers. The simple input-to-output network that results is mathematically equivalent to multivariate multiple linear regression.

By representing words' meanings numerically, it becomes possible to harness the power of distributional semantics (Landauer & Dumais, 1997; Mikolov et al., 2013; Mitchell & Lapata, 2008) when considering the questions of what possible meanings are given words' forms, and what possible forms are given words' meanings. This is important, because form and meaning can show intricate interactions. For instance, Baayen and Moscoso del Prado Martín (2005) called attention to irregular verbs in English (and as well in German and Dutch) being more similar to each other in their meanings than regular verbs. The greater semantic density of irregular verbs in English may underlie the interaction of semantic deficits and regularity in aphasia reported by Bird et al. (2003), and modeled computationally using distributional semantics by Heitmeier and Baayen (2021). Below, we shall see that the broken plurals and the sound plurals of Maltese may also pattern differently in semantic space.

Several studies suggest that the DL correctly predicts the forms of complex words (see Baayen et al. (2018) for Latin verb inflection, Chuang et al. (2020) for Estonian noun inflection, van de Vijver and Uwambayinema (accepted) and van de Vijver et al. (2021) for Kinyarwanda nouns and verbs, and Chuang, Kang, Luo, et al. (2021) for Korean verbs). The first goal of the present study is to clarify whether the theory of the DL also correctly predicts Maltese singular and plural nouns. Of particular interest is how well the simple networks used by the DL are able to model not only concatenative morphology, but also non-concatenative morphology.

The framework of the DL has also been used to predict how words are realized phonetically. Tomaschek et al. (2021) modeled the duration of English word-final [s] for different grammatical functions, Saito et al. (2021) used measures from the model to predict tongue trajectories, Chuang, Vollmer, et al. (2021) predicted word duration for English pseudowords as pronounced by native speakers of English, and Chuang, Kang, Luo, et al. (2021) applied the model to word duration in Taiwan Mandarin. The latter study also shows that the priming effects reported for Dutch in Creemers et al. (2020) are correctly predicted by the model (see also Baayen & Smolka, 2020, for German). In the light of these results, the second goal of the present study is to clarify whether measures derived from the model help predict lexical processing costs, as gauged with a cross-modal primed lexical decision task.

The remainder of this paper is structured as follows. We first provide an overview of plural formation in Maltese and report previous experimental and computational studies on Maltese plurals. Section 3 proceeds with an introduction to the 'Discriminative Lexicon'. We then present the computational models that we developed for the Maltese noun system. We report how well they perform as a memory for known words, and also examine the extent to which the memory is productive, in the sense that it can handle unseen words that it has not been trained on. Subsequently, we show how the theory can be used to obtain further insight into the lexical processing of Maltese nouns in comprehension. We conclude this study with a discussion of the new insights that our results bring to morphological theory on the one hand, and its limitations on the other hand.

2 Maltese plurals

The turbulent history of Malta is reflected in the national language of the island. Maltese developed from Maghrebi Arabic, and has absorbed influences from Sicilian, Italian and, more recently, from English. These influences affected its lexicon and its morphology (Hoberman, 2007).

The Maltese noun plural system shows a perplexing amount of possible plural forms. Maltese has a great number of typically Semitic non-concatenative plural forms—called broken plurals in the Semitic linguistic tradition. Broken plurals are characterized by differences in the prosodic structure of a plural as compared to its corresponding singular form. For example, the singular form *kelb* 'dog' /kɛlp/ has the plural form *klieb* 'dogs' /kliːp/[2] in which the coda consonant [l] of the singular is found in the onset

[2]Another possible phonetic variant given in the online dictionary Ġabra is /klɪːp/

of the plural form. In addition, the vowel [ɛ] in the singular form corresponds to [iː] in the plural. Schembri (2012) distinguishes 11 different broken plural patterns. In Maltese, broken plurals account only for a small proportion of plural forms of the language (Borg & Azzopardi-Alexander, 1997, report a proportion of 10%). In addition to broken plurals, Maltese also has a sizable set of sound plurals and the majority of plurals belong to this category (Borg & Azzopardi-Alexander, 1997; Nieder, van de Vijver, et al., 2021a).

Sound plurals are characterized by additional segmental material at the right side of the plural in comparison to the singular: The singular form *prezz* 'price' has the plural form *prezzijiet* in which the plural differs from the singular due to the presence of a particular plural exponent, the suffix -*ijiet*. In their work, Nieder, van de Vijver, et al. (2021a, 2021b) distinguish 12 different sound plural patterns (they count the dual forms as a sound plural pattern) with different frequency distributions and productivity. Table 1 below gives an overview of the Maltese sound and broken plural patterns and the two possible dual forms.

The complexity of the Maltese noun system stems from two sources. One is the sheer variety of suffixes and patterns exhibited in plurals. This sets Maltese apart from languages in which the complexity of nominal systems is due to nouns falling into different declension classes. The other complexity is the availability of several plural forms for many singulars, without there being a noticeable semantic difference among the plural variants. For example, the singular *kaxxa* (sg.) 'box' has two plural forms, one is a broken plural, *kaxex*, and one is a sound plural, *kaxxi*; another example is the singular *giddieb* (sg.) 'liar', which has two sound plural forms, *giddieba* and *giddibin*.

In addition to sound and broken plurals, Maltese shows other plural types for a small number of nouns, such as the suppletive plural, e.g. *mara* - *nisa* 'women' or a double plural marking that is a blend of a broken plural and a sound plural suffix (called *plural of the plural* by Mayer et al. (2013)), the singular *tarf* has the blended plural *trufijiet* 'edge'. A few words are pluralized with a dual suffix but grammatically behave like plural words, for example *sieq* - *saqajn* 'foot' (Borg & Azzopardi-Alexander, 1997; Mayer et al., 2013).

2.1 Experimental and computational research on Maltese plurals

There exists both experimental and computational research on the Maltese nominal system. In the following, we first discuss the experimental research on Maltese nouns before turning to the computational studies.

Singular	Plural	Gloss	Plural Type
fardal	*fra:dal*	‘aprons’	broken A, CCVVCVC
birra	*birer*	‘beers’	broken B, (C)CVCVC
kbir	*kba:r*	‘big (pl.)’	broken C, CCVVC
ftira	*ftajjar*	‘type of bread (pl.)’	broken D, CCVjjVC
bitħa	*btieħi*	‘yards’	broken E, CCVVCV
sider	*isdra*	‘chests’	broken F, VCCCV
marid	*morda*	‘sick persons’	broken G, CVCCV
għodda	*għodod*	‘tools’	broken H, (għ)VCVC
elf	*eluf*	‘thousands’	broken I, VCVC
għaref	*għorrief*	‘wise men’	broken J, CVCCVVC(V)
għama	*għomja*	‘blind persons’	broken K, (għ)VCCV
karta	*karti*	‘paper’	sound, -*i*
omm	*ommijiet*	‘mother’	sound, -*ijiet*
rixa	*rixiet*	‘feather’	sound, -*iet*
giddieb	*giddieba*	‘liar’	sound, -*a*
meħlus	*meħlusin*	‘freed’	sound, -*in*
kuxin	*kuxins*	‘cushion’	sound, -*s*
triq	*triqat*	‘street’	sound, -*at*
sid	*sidien*	‘owner’	sound, -*ien*
baħri	*baħrin*	‘sailor’	sound, -*n*
ħati	*ħatjin*	‘guilty’	sound, -*jin*
qiegħ	*qigħan*	‘bottom’	sound, -*an*
spalla	*spallejn*	‘shoulder’	dual, -*ejn/ajn*
sieq	*saqajn*	‘foot’	dual, -*ejn/ajn*

Table 1: Maltese broken plurals, sound plurals and duals (examples taken from Nieder, van de Vijver, et al., 2021a; Schembri, 2012). The words are provided in Maltese orthography, which is a close approximation of a broad phonetic transcription, except in two cases. First, long a is not indicated in orthography. We therefore added a colon to long a. Second, the digraph għ is historically a pharyngeal fricative, which was lost in modern Maltese (Borg & Azzopardi-Alexander, 1997).

Two experimental studies have clarified that native speakers use information about pattern frequency to produce and process plural forms for singulars they never heard before (Nieder, van de Vijver, et al., 2021a, 2021b). While some plural suffixes and patterns occur frequently in the language, for example the sound plural forms ending in *-i* and *-ijiet* or the broken plural patterns characterized by the CV-templates $CCVVCVC$ (broken A) and $CCVVC$ (broken C), others are found in a relatively small number of plural forms only (see Nieder, van de Vijver, et al., 2021a, 2021b; Schembri, 2012, for detailed information about pattern frequency in Maltese).

In a production study, Nieder, van de Vijver, et al. (2021a) asked Maltese native speakers to produce plurals for existing singulars and pseudo-singulars. The plurals produced by the participants reflected the frequency of the plural patterns in Maltese. The participants made use of more frequent plural suffixes when they produced sound plurals and of more frequent CV templates when they produced broken plurals (a finding that is also reported by Drake (2018) for Maltese diminutives).

Further evidence for the importance of the type frequency of exponents (sound plurals) and CV templates (broken plurals) emerged from a reaction time study by Nieder, van de Vijver, et al. (2021b). Frequent broken templates and frequent sound plural exponents elicited significantly shorter reaction times than infrequent ones. This experiment did not provide evidence for an effect of plural type (broken versus sound): on average, response times for both kinds of plurals were highly similar. Below, we return to this study, to show that nevertheless the way in which responses are generated in this task differs for broken plurals and sound plurals.

Computational analyses of the Maltese plural formation have focused on form-to-form modeling using sets of rules or using analogical mappings. These computational studies are moving away from an earlier consensus among Maltese scholars, according to which there are no rules governing broken plurals (as discussed in Schembri, 2012). Invariably, the singular form is taken as starting point for predicting the corresponding plural form. Some models are classifiers for plural classes, others generate full plural forms given the corresponding singulars.

Mayer et al. (2013) present a computational study of Maltese broken plurals that focuses on the application of rules to form plurals from singulars. They propose a set of four rules, based on the work of Schembri (2012), which derives broken plurals from their singulars. These rules were shown to correctly derive 75% of all 654 forms in their database that have a broken plural. This study shows unambiguously that the Maltese broken plurals are to a considerable extent systematic, but it does not address the question of how speakers select between broken and sound plurals. Furthermore, as mentioned above, it is not self-evident from a cognitive perspective that speakers would create plurals from singulars.

Farrugia and Rosner (2008) also focused exclusively on broken plurals, using an artificial neural network with encoder and decoder hidden layers, to categorize and produce Maltese broken plurals. As basis for their work they also edited the analysis of Schembri (2012). Operating on phoneme-based representations, their model categorized nearly all nouns in their dataset with an accuracy of around 98%. Although they report good results for forms the model had seen in training, it did not perform well on unseen forms, achieving exact matches between predicted and observed plural forms for only 26.6% of the cases. This computational model again shows that there are indeed systematic relations between the form of the singular and its broken plural form, and that these relations can be derived from the data without requiring handcrafted rules. It remains unclear, however, how the model would have performed if it had been trained on both broken plurals and sound plurals jointly.

Nieder, Tomaschek, et al. (2021) compared three different computational models to investigate whether it is in principle possible to account for the form-based relations in Maltese nominal paradigms without taking recourse to the construct of the morpheme: the Tilburg Memory-Based Learner (TiMBL) (Daelemans et al., 2004), the Naive Discriminative Learner (NDL) (Baayen, 2011), and an Encoder-Decoder network. TiMBL and NDL are classifiers, the Encoder-Decoder network is a model generating actual plural forms. Models were trained on a dataset consisting of both sound plurals and broken plurals. The classifiers were given the task to predict which class out of 8 plural classes (4 broken plural classes, and 4 sound plural classes: three for the three most frequent exponents, and one for all other exponents) is appropriate for a given singular. TiMBL's best performance under 10-fold cross-validation was 97%, whereas NDL's best performance under 10-fold cross-validation was 88.7%. The best performance of the Encoder-Decoder model was at 48.22%. Interestingly, although information about the CV template has been reported to increase classification accuracy for Arabic (Dawdy-Hesterberg & Pierrehumbert, 2014), such information did not improve the accuracy of the TiMBL classifier for Maltese.

What all these modeling studies clarify is that there is considerable structure in the Maltese noun system. However, the best-performing models are either trained on only broken plurals, or they are trained to predict form classes, including classes that lump together less frequent form changes. Furthermore, all models focus on production, predicting plurals from singulars without considering words' meanings, and do not address the comprehension of Maltese nouns. In what follows, we address this broader range of questions within the framework of the Discriminative Lexicon. Before doing so, we first introduce the dataset that we used for training and evaluating our models.

2.2 Dataset

The dataset consists of all broken plurals listed by Schembri (2012) and all word forms tagged as nouns from the MLRS Korpus Malti version 2.0 and 3.0 (Gatt & Čéplö, 2013). The resulting list of nouns was then enriched with information extracted from a Maltese online dictionary (*Ġabra*, Camilleri, 2013) using the free corpus tool *Coquery* (Kunter, 2017), resulting in a dataset with singulars, their corresponding plurals and their glosses. Subsequently, the dataset was manually extended with information about grammatical number (broken vs. sound plural, dual or suppletive), CV structure, number of occurrences (based on the Korpus Malti v. 2.0 and 3.0), origin (Semitic vs. Non-Semitic), grammatical gender (based on Aquilina (1987)), concreteness (abstract vs. concrete), and type of noun (verbal noun or collective noun).

The resulting dataset contains 6511 word forms in total: 3364 plurals, 3132 singulars and 15 dual forms. Of the 3364 plurals, 892 are broken plural forms while 2458 are sound plural forms (with a total of 11 different sound plural types and 11 different broken plural types), reflecting the proportion of plural types in use in Maltese. The remaining 29 nouns of our dataset labeled as plurals have plurals that are neither of the broken nor of the sound type: 8 of these words have a double plural marking, e.g. *sema* (sg.) - *smewwiet* (pl.) 'sky', which is a combination of a broken and a sound plural. Fifteen words are dual forms, such as *id* (sg.) - *idejn* (dual) 'hands', and 6 words have a suppletive plural, e.g. *mara* (sg.) - *nisa* (pl.) 'women', see Borg and Azzopardi-Alexander (1997) for further details.

3 Predicting Maltese noun inflection

The models for the Maltese plurals reviewed in section 2.1 all predict the appropriate form of a plural from its corresponding singular. However useful rules for building forms from other forms may be for the teaching of a second language, it is far from clear that native speakers and young L1 learners would follow the same procedure (Blevins, 2016; Dell, 1986; Levelt et al., 1999; Zamuner et al., 2011). The DL model proposed by Baayen et al. (2019) takes as its point of departure that the task of morphology is to explain how listeners understand complex words, and how speakers produce them. In other words, the DL focuses on understanding words' meanings given their forms, and producing words' forms given their meanings. Furthermore, the relation between form and meaning is modeled as immediate, without any further intervening layers of representations.

The central ideas underlying the perspective of DL on form and meaning are illustrated in Figure 1. In the upper left, the matrix $\boldsymbol{C}$ specifies, for three words w_1, w_2, w_3, their respective form vectors with values for two form features, f_1 and f_2. In the upper right, the matrix $\boldsymbol{S}$ specifies the semantic vectors for the same three words, which have values on the semantic dimensions s_1 and s_2. The form vectors are displayed in the lower left, and the semantic vectors in the lower right. The mapping $\boldsymbol{F}$ takes the red vectors and changes them into the blue vectors. Formally, this is done by post-multiplying $\boldsymbol{C}$ with $\boldsymbol{F}$: $\boldsymbol{CF} = \boldsymbol{S}$. Conversely, the $\boldsymbol{G}$ matrix takes the blue vectors and maps them onto the red vectors: $\boldsymbol{SG} = \boldsymbol{C}$. The mappings that the DL sets up between numeric vectors representing forms and numeric vectors representing meanings are the simplest mappings possible. They can be conceptualized as simple artificial neural networks connecting form units (f_1, f_2) and semantic units (s_1, s_2). In other words, the mappings implement full connectivity between all form units and all semantic units. The networks do not make use of any hidden layers. Equivalently, the mappings of the DL can also be understood as implementing multivariate multiple regression. For comprehension, for instance, the $\boldsymbol{F}$ matrix can be interpreted as the matrix with beta coefficients of a regression model. The beta weights in the first column of $\boldsymbol{F}$ are used to predict the response variable given by the first column of $\boldsymbol{S}$. Likewise, the beta weights in the second column of $\boldsymbol{F}$ are used to predict the response variable given in the second column of $\boldsymbol{S}$. The same logic applies to the beta weights in $\boldsymbol{G}$: For instance, the beta weights in the first column are used to predict the response variable in the first column in $\boldsymbol{C}$. The method that we used to estimate the mappings $\boldsymbol{F}$ and $\boldsymbol{G}$ is taken from linear algebra, for technical details, the reader is referred to Shafaei-Bajestan et al. (2021).

In general, for a given set of n words and m dimensions in which differences in form are expressed, we bring together their numeric form vectors into an $n \times m$ form matrix $\boldsymbol{C}$. Given k-dimensional vectors representing words' meanings, we set up an $n \times k$ semantic matrix $\boldsymbol{S}$. The $m \times k$ mapping $\boldsymbol{F}$ takes the vectors in $\boldsymbol{C}$ and transforms these vectors as precisely as possible into the semantic vectors of $\boldsymbol{S}$. This is accomplished by solving the equation $\boldsymbol{CF} = \boldsymbol{S}$. For production, the DL model posits a $k \times m$ mapping $\boldsymbol{G}$ from the meaning vectors $\boldsymbol{S}$ to the form vectors in $\boldsymbol{C}$. This matrix is estimated by solving $\boldsymbol{SG} = \boldsymbol{C}$. For all but the smallest toy examples, the predicted form vectors $\hat{\boldsymbol{C}} = \boldsymbol{SG}$ will only approximate the targeted gold-standard form vectors $\boldsymbol{C}$, which is why, following statistical practice, we use the notation $\hat{\boldsymbol{C}}$

$$
\boldsymbol{C} = \begin{array}{c} \\ w_1 \\ w_2 \\ w_3 \end{array} \begin{array}{c} \begin{array}{cc} f_1 & f_2 \end{array} \\ \begin{pmatrix} 1 & 2 \\ -2 & -2 \\ -2 & 1 \end{pmatrix} \end{array} \qquad \boldsymbol{S} = \begin{array}{c} \\ w_1 \\ w_2 \\ w_3 \end{array} \begin{array}{c} \begin{array}{cc} s_1 & s_2 \end{array} \\ \begin{pmatrix} 2 & -4 \\ -4 & 4 \\ -4 & -2 \end{pmatrix} \end{array}
$$

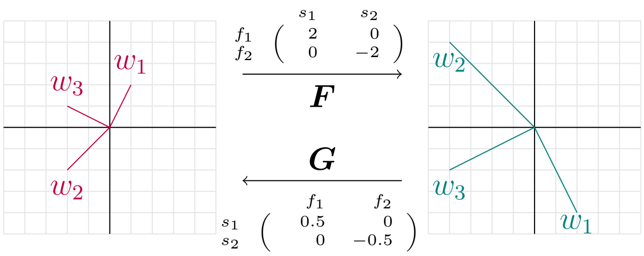

Figure 1: Linear mappings between form vectors (the row vectors of $\boldsymbol{C}$, displayed in red on the left) and meaning vectors (the row vectors of $\boldsymbol{S}$, displayed in blue on the right). The mapping $\boldsymbol{F}$ changes the form vectors into semantic vectors, and the inverse mapping $\boldsymbol{G}$ takes the semantic vectors and changes them into the form vectors. The mappings $\boldsymbol{F}$ and $\boldsymbol{G}$ define networks, the weights on connections from form features f to semantic features s, and from semantic features s to form features f are given by the respective entries in the mapping matrices.

rather than $\boldsymbol{C}$. The same holds for the predicted semantic vectors $\hat{\boldsymbol{S}}$. Nevertheless, the estimated weights are optimal, in the sense that they minimize the mean squared error. They represent the 'endstate' of learning that a simple two-layer artificial neural network can achieve by endlessly iterating through the training data with the incremental learning rule of Widrow and Hoff (1960). In what follows, we refer to the learning of the mappings using the mathematics of multivariate linear regression as 'Linear Discriminative Learning' (LDL).

3.1 Constructing the form matrix

	Lexeme	Number	Gender
kelb	KELB	singular	M
kelba	KELB	singular	F
klieb	KELB	plural	M, F

Table 2: Paradigm for the Maltese noun *kelb* 'dog'.

To illustrate the central concepts of LDL, consider the Maltese toy lexicon listed in Table 2. This lexicon consists of a singular word for a male dog, a singular word for a female dog and the plural word for both.

The first modeling step is to make a decision as to how these word forms can be represented as numeric vectors. One possibility is to decompose word forms into triphones, which target, in a crude way, context-sensitive phone representations. Heitmeier et al. (2021) present a systematic overview of modeling options for word form representations in LDL. They report best generalizations for triphones (as compared to biphones or quadrophones) due to their discriminatory power as a result of a balanced number of unique cues (see Heitmeier et al., 2021). For our example lexicon, there are 11 distinct triphones. We couple each distinct triphone with a form dimension. Words that contain a given triphone receive the value 1 for this dimension, and otherwise the value 0. For our example lexicon, we obtain the following form matrix $\boldsymbol{C}$:

$$\boldsymbol{C} = \begin{array}{l} \\ kelb \\ kelba \\ klieb \end{array} \begin{array}{c} \begin{array}{ccccccccccc} \#ke & kel & elb & lb\# & lba & ba\# & \#kl & kli & lie & ieb & eb\# \end{array} \\ \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \end{array}$$

In this form matrix, the hash mark # represents a word boundary.

Instead of representing words' forms by indicating which triphones are present, we can set up form vectors that decompose a word's form into its constituent syllables. In this study, again based on the results of Heitmeier et al. (2021), we opted for bi-syllable cues that are not only a linguistically-informed unit driving articulation (Levelt et al., 1999) but are also known to capture certain suprasegmental effects (Heitmeier et al., 2021). Below, we report results for simulations using these two ways of representing word form information.

3.2 Constructing the semantic matrix

The row vectors of the semantic matrix $\boldsymbol{S}$ represent a word form's meaning numerically. Within the general framework of distributional semantics, many algorithms are now available for deriving semantic vectors (known as embeddings in computational linguistics) from corpora (Baroni et al., 2014; Bojanowski et al., 2017; Joulin, Grave, Bojanowski, Douze, et al., 2016; Joulin, Grave, Bojanowski, & Mikolov, 2016; Mikolov et al., 2013; Pennington et al., 2014; Yang et al., 2017). In the present study, we explore two kinds of semantic vectors: vectors that we constructed ourselves in a linguistically informed way, which we call simulated vectors, and ready-made vectors that were generated with `fasttext` (Joulin, Grave, Bojanowski, Douze, et al., 2016; Joulin, Grave, Bojanowski, & Mikolov, 2016), which we call corpus-based vectors.

3.2.1 Simulated vectors

The row vectors of the semantic matrix $\boldsymbol{S}$ represent words' meanings in a high-dimensional space. We can simulate such vectors using a random number generator. The idea underlying this approach is similar to the statistical concept of 'generating' a statistical model: when we model a response variable y as a linear function of x,

$$y_i = a + bx_i + \epsilon_i,$$

the hope is that we can generate a dataset that has all the properties of the observed data, with as only difference the measurement errors ϵ_i. When simulating semantic vectors, we do the same: we set up a model that generates semantic vectors that represent the semantic structure of words, apart from word-specific or idiosyncratic aspects of words' meanings (see, e.g., Booij, 1996; Sinclair, 1991a, for word-specific semantics of inherent inflection). For our example lexicon, we generated 11-dimensional vectors, matching the dimensionality of the form matrix $\boldsymbol{C}$. The result is a straightforward table with real-valued numbers:

$$\boldsymbol{S} = \begin{array}{c} \\ \textit{kelb} \\ \textit{kelba} \\ \textit{klieb} \end{array} \begin{array}{c} \begin{array}{rrrrrrrrrrr} S1 & S2 & S3 & S4 & S5 & S6 & S7 & S8 & S9 & S10 & S11 \end{array} \\ \begin{pmatrix} 0.46 & 4.16 & 8.50 & -4.46 & 8.96 & -4.11 & 8.42 & 9.21 & -25.75 & 15.83 & -14.93 \\ 0.61 & -11.93 & 8.09 & 1.00 & 3.44 & -11.98 & 8.72 & -4.75 & -33.29 & 10.39 & -2.12 \\ 5.67 & 9.84 & 11.26 & 0.85 & 10.69 & -4.24 & 0.21 & 4.81 & -26.47 & 10.82 & -11.76 \end{pmatrix} \end{array}$$

However, if we use this method to create semantic vectors for each word form, then, unavoidably, the resulting semantic vectors are almost completely uncorrelated, which implies that the meanings of these words are understood to be semantically entirely unrelated. When considering monomorphemic words, such uncorrelated vectors are justifiable as a very first approximation that is no worse (but also no better) than representing words' meanings by their own symbolic nodes. However, since inflected words share inflectional features, we need to generate vectors that properly reflect that for instance plurals are semantically more similar to other plurals, and less similar in meaning than singulars.

Following Baayen et al. (2019), we generated semantic vectors of inflected words by taking the (generated) vector of the lexeme and adding to it additional (generated) vectors, one for each inflectional function. For the Latin noun *horti* ('garden', genitive singular), for instance, a vector for genitive and a vector for singular are added to the vector of GARDEN:

$$\overrightarrow{\text{horti}} = \overrightarrow{\text{GARDEN}} + \overrightarrow{\text{SINGULAR}} + \overrightarrow{\text{GENITIVE}}.$$

For Maltese nouns, we considered several semantic features: whether a noun is derived from a verb (e.g., participles), whether a noun has collective semantics, whether a noun has masculine or feminine gender, and number. The former two features were coded as privative oppositions, i.e., we added a vector representing collective semantics to collective meanings, but left the semantic vectors of all other nouns unchanged. For the latter two features, we generated semantic vectors under the assumption that here we have equipollent oppositions. For number, we thus decided to construct three semantic vectors, one for singular meaning, one for dual meaning, and one for plural meaning. For the forms *kelb, kelba* and *klieb*, the semantic vectors in our example lexicon given above (matrix $\boldsymbol{S}$) were obtained by adding the pertinent inflectional vectors to the vectors of the lexemes, together with error vectors representing words' semantic idiosyncracies:

$$\begin{array}{llllllll} \textit{kelb}: & \overrightarrow{\text{KELB}} & + & \overrightarrow{\text{SINGULAR}} & + & \overrightarrow{\text{MASCULINE}} & + & \overrightarrow{\varepsilon} \\ \textit{kelba}: & \overrightarrow{\text{KELB}} & + & \overrightarrow{\text{SINGULAR}} & + & \overrightarrow{\text{FEMININE}} & + & \overrightarrow{\varepsilon} \\ \textit{klieb}: & \overrightarrow{\text{KELB}} & + & \overrightarrow{\text{PLURAL}} & + & \overrightarrow{\text{MASCULINE}} & + & \overrightarrow{\varepsilon} \end{array}$$

An alternative coding for number, that we did not pursue, would be to code number as a privative opposition, with an unmarked singular and marked dual and plural. However, as the broken plurals are formally not marked variants of their corresponding singulars, we opted for implementing equipollent semantic vectors for number.

In summary, we generate semantic vectors for inflected forms by addition of the primitive vectors for their constituent meanings. This additive process is the way in which we approximate the conceptualization of the semantics of inflected words.

3.2.2 Corpus-based vectors using `fasttext`

Although simulated vectors have been found useful for modeling morphological processing in comprehension and production, they make the simplifying assumption that all base word lexemes are semantically unrelated: their simulated semantic vectors are almost completely orthogonal. In addition, the way in which inflectional semantics is accounted for may also require more precision, see, e.g., Shafaei-Bajestan et al. (2022) for discussion of the semantics of the English noun plural. Instead of working with simulated vectors, Baayen et al. (2019) derived semantic vectors for both content lexemes and inflectional functions such as singular and plural by first morphologically tagging a corpus (in their study, the TASA corpus, Ivens & Koslin, 1991), and then using a method from distributional semantics to construct semantic vectors for both content words and for the inflectional (as well as derivational) functions identified by the tagger.

Since computational resources for Maltese are limited, for the present study, we complemented modeling using simulated vectors with modeling using ready-made vectors that were created with `fasttext` (Joulin, Grave, Bojanowski, Douze, et al., 2016; Joulin, Grave, Bojanowski, & Mikolov, 2016). `Fasttext` is an open-source library for text classification and representation that offers the possibility to train a `fasttext` model on a set of data or to download pre-trained vectors for various languages from https://fasttext.cc/docs/en/crawl-vectors.html. For this study, we opted for the latter approach.

Modeling with `fasttext` vectors has as advantage, compared to simulated vectors, that the LDL mappings will be able to take into account similarities in meaning between content words, as well as inflectional similarities. However, the algorithm underlying `fasttext` constructs semantic vectors for words from semantic vectors of substrings of words by representing words as a sum of their character n-grams (see Joulin, Grave, Bojanowski, Douze, et al., 2016; Joulin, Grave, Bojanowski, & Mikolov, 2016, for details on how the vectors were created). As a consequence, it cannot be completely ruled out that for inflected words the algorithm is capturing not only similarities in meaning but also similarities in form.

We extracted `fasttext` vectors for the word forms in our data set using the pre-trained 300 dimensional word-vectors that are available for Maltese at https://fasttext.cc/docs/en/crawl-vectors.html. For 4056 of the 6511 nouns in our dataset, `fasttext` vectors were available; of these 4056 word forms, 2266 are singulars and 1781 are plurals.

In order to obtain some insight in how well `fasttext` captures the difference between singular and plural meaning, we projected the 300-dimensional `fasttext` space onto a 2-dimensional plane using Principal Components Analysis. A scatterplot of nouns in the plane of the first two principal components, color-coded for number and plural type, is shown in Figure 2. Interestingly, we find distinguishable clusters of singulars (light green) and plurals (orange, dark green), albeit with considerable overlap. In addition, sound plurals (orange) and broken plurals (dark green) seem to dwell in somewhat different semantic subspaces as well. This is confirmed by a Linear Discriminant Analysis (LDA), which showed that a classification of singular, sound plural and broken plural words using the first fifty principal components reaches 85% classification accuracy. Apparently, number and type of plural are to some extent intertwined with word meaning. This interaction of regularity with semantics replicates a similar interaction for English regular and irregular verbs reported by Baayen and Moscoso del Prado Martín (2005).

Figure 3 addresses how well `fasttext` captures differences in gender. Despite substantial overlap of the clusters, Linear Discriminant Analysis, again using the first fifty principal components, achieved a classification accuracy of 79% and 70% for singular and plural words respectively. For the other semantic features labeled in our dataset (concreteness, verbal noun, collective noun), however, due to the fact that usually one level has overwhelmingly more tokens than the other, no clustering in the semantic space could be observed.

Above, we mentioned that `fasttext` looks "into" words by representing word forms as a bag of n-grams, and that as a consequence, it cannot be ruled out a-priori that similarities in meaning are confounded with similarities in form. However, given the complexities of the Maltese plural forms, it is unlikely that the clustering visible in Figure 2 is driven predominantly by form similarity. Nevertheless, replication of this interaction of plural type and semantics using, for instance, `word2vec` (Mikolov et al., 2013), would strengthen the present conclusions for Maltese.

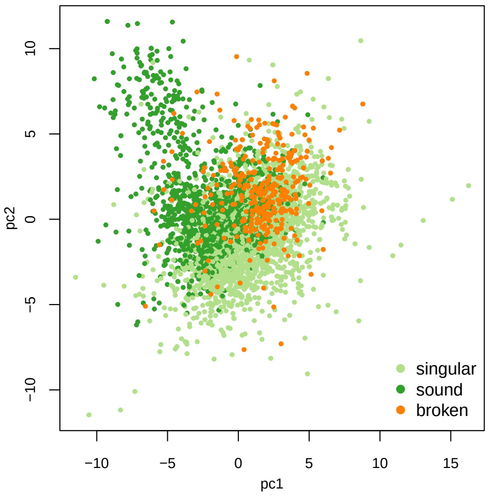

Figure 2: Projection of `fasttext` semantic vectors onto a two-dimensional plane. Number and plural types (sound and broken) are color-coded. Singulars and broken plurals cluster more to the right on PC1, whereas sound plurals and broken plurals cluster more to the top on PC2.

3.2.3 Evaluating model performance

Before reporting how well the DL model approximates the Maltese noun system, we need to explain how we evaluate model performance.

To evaluate comprehension, we calculated the correlations between a given word's predicted semantic vector ($\hat{s}_i$) and all the gold standard semantic vectors in the lexicon (the row vectors of $\boldsymbol{S}$). If $\hat{s}_i$ has the highest correlation with the semantic vector of the targeted word (s_i), comprehension is considered successful. On the other hand, unsuccessful comprehension occurs when the highest correlation is with another word than target word. It should be noted that for homophones, we consider comprehension correct as long as $\hat{s}_i$ has the best correlation with one of the homophone meanings, e.g. Maltese *xark* 'shark' /ʃɐrk/ and *xark* /ʃɐrk/ 'a person who conducts business shrewdly or acts for their own material benefit' (note that there also is a Semitic word to express 'shark' available in Maltese: *kelb il-baħar*). This is because here we are modeling the processing of words in isolation. Given that it is not possible to recognize a specific homophone meaning out of context, we therefore adopted this lenient evaluation metric for comprehension.

With respect to production, as a first step, we generated for each word i the predicted form vector $\hat{c}_i$ from its semantic vector s_i. This predicted form vector, however, only provides information about the amount of semantic

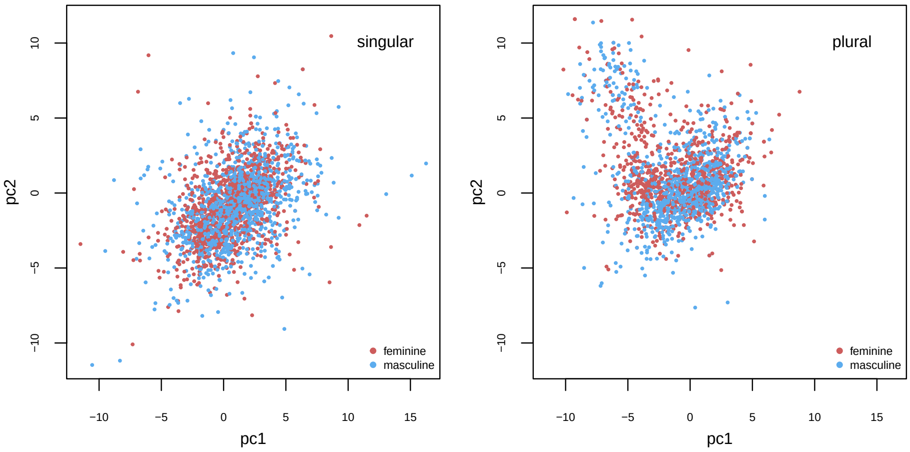

Figure 3: Projection of `fasttext` semantic vectors onto a two-dimensional plane, spanned by the first two principal components. The left panel plots singular feminine (red) and masculine (blue) words, and the right panel plots plural words. PC2 captures to some extent plurality, whereas PC1 captures aspects of gender, resulting in somewhat differentiated clustering within number for feminine vs masculine nouns.

support for the sublexical cues (such as triphones or bi-syllables); it does not inform us about the order in which well-supported cues have to be placed for articulation. For ordering, the model makes use of the order information that is implicit in the sublexical cues. Take triphones, for example. The triphone `kel` can be followed by `elb` (to form the word *kelb*), given the identity of the final diphone `el` in `kel` and the initial diphone `el` in `elb`. In the absence of such overlap (e.g., for `kel` and `lie`), no sequential ordering is possible. As the lexicon becomes larger, the number of possible triphone combinations also grows, resulting in multiple candidate forms for a given form vector $\hat{c}_i$. The candidate selected for articulation is chosen such that it best realizes the meaning the speaker has in mind. Technically, this is accomplished by first generating, for each candidate form ω_j its predicted semantic vector $\hat{s}_j$, and then selecting from these semantic vectors the one that is most similar to the targeted semantic vector $\boldsymbol{s}_i$ that is to be expressed. In other words, we generate the predicted semantic vectors for all candidate forms and select as model prediction the form vector associated with the predicted semantic vector that has the closest meaning to the targeted meaning (a process called 'synthesis-by-analysis' by Baayen et al., 2018).

For the simulations presented in this study, we used the **JudiLing** package, an implementation of LDL in the `Julia` language (Luo et al., 2021).

For production, we used the `learn_paths` function for ordering sublexical features into words. This algorithm takes predicted form vectors, and learns to predict at what position(s) in a word a sublexical cue occurs. In this way, each of a word's sublexical cues is associated with a number reflecting how well it is supported for its position in the word. We refer to this number as a cue's positional support. Only cues with sufficient positional support are taken into account when assembling the set of word candidates. What counts as sufficient positional support is determined by a threshold value θ: Only words with a positional support exceeding θ are taken into consideration. More detail about the `learn_paths` algorithm can be found in Luo et al. (2021). In Section 5.4, we will show that the total amount of positional support for a word's cues is predictive for reaction times to Maltese plurals.

4 Modeling results

4.1 Evaluation on training data

With two cue representations (one using triphones and one using bi-syllables) and two semantic representations (one using simulated semantic vectors and the other using `fasttext` vectors), we have in total four models. For comparison, the dimension of the simulated vectors is set to 300, mirroring the dimensionality of the `fasttext` vectors. It should be noted, however, that since `fasttext` vectors are not available for all the word forms in the dataset, we therefore worked with a smaller dataset ($n = 4056$) when using `fasttext`. Comprehension and production accuracies of the four models are presented in Table 3. For comprehension, bi-syllable as cues yielded higher accuracies than triphones as cues, regardless of the kind of semantic representation. With respect to production, we again see an overall advantage of bi-syllable cues. In addition, while the difference between vector types is not large for bi-syllable cues, with triphone cues, model performance with simulated vectors is a lot worse than that with `fasttext` vectors. Given that there are more bi-syllable cues than triphone cues (9821 *vs.* 4272 for the full dataset), the absence of sufficient semantic structure in simulated vectors (e.g., all lexeme and inflectional features are orthogonal to one another) seems to be potentially harmful when the form space is not well differentiated. On the other hand, the model may be overfitting when bi-syllable cues are used. We return to this possibility below.

Given the high accuracy of all four models, we can conclude that the model generally has a good memory for understanding and producing Maltese plurals. However, we do not know how the model performs with respect to inflected forms that it has not encountered during training. In other words, we do not yet know to what extent the model is productive. To test this, we ran the model on a subset of the data that was not used during training. The results of this process are reported in the following section.

	comprehension		production	
	simulated	fasttext	simulated	fasttext
triphone	93.1%	95.6%	82.9%	91.1%
bi-syllable	99.8%	99.9%	99.0%	96.3%

Table 3: Model performance for comprehension (left) and production (right) for four combinations of cue and semantic representations. For production, the threshold was set to 0.005 for both the simulated and the `fasttext` models.

4.2 Evaluation on held-out data

The question of whether our model is productive for Maltese is of considerable theoretical interest because the noun system of Maltese is not straightforwardly regular. Although some rules can be formulated, indicating that the system is not just random, the many patterns for broken plurals and the wide variety of plural exponents characterize a system for which full productivity cannot be expected. It would actually be strange and worrisome if computational models were to be able to predict unseen forms with close to 100% accuracy. Since regularity is generally seen as a prerequisite for productivity, the Maltese noun system is perhaps best characterized as semi-productive. This possibility receives support from the observation that native speakers of Maltese are often unsure about what the proper plural of an unknown or infrequent word might be, as indicated by the production study in Nieder, van de Vijver, et al. (2021a). In the light of these considerations, a substantial drop in prediction accuracy is expected for held-out data, compared to the accuracy for the training data.

We also expect to find that for held-out data, production accuracy will be somewhat lower than comprehension accuracy. This is due to the 'synthesis-by-analysis' approach of the model: to select a candidate path for production, the LDL model is using the results from the comprehension model (see Baayen et al., 2018; Heitmeier et al., 2021, for further details). The familiar asymmetry between production and comprehension (Boersma, 1998; Pater, 2004; Smolensky, 1996) was already visible in the results for the training data (see Table 3), and we anticipate it will be present, and perhaps more pronounced, for the held-out data.

To examine model productivity, we held out 10% of the words in our dataset as testing data. The held-out words were selected based on the criterion that all the sublexical cues and inflectional features of the words have already been available to the model during training. Furthermore, the held-out words were constrained to have lexemes that occurred in the training data. In addition, we used the smaller dataset, instead of the full dataset, to enable comparisons between simulated and `fasttext` vectors.

The testing data contained 174 singular forms and 205 plural forms. Of the plural forms, 192 were sound plurals, 12 were broken plurals, and one was a suppletive form. During training, bi-syllable cues consistently outperformed triphone cues, both with simulated semantic vectors and with `fasttext` vectors. This indicates that the models with bi-syllable cues, which outnumber triphone cues, are not overfitting. In what follows, we only report results obtained with bi-syllables.[3]

For comprehension, simulated vectors performed better than `fasttext`, with an accuracy at 77.8% and 63.6% respectively. When the top 10 candidate meanings are considered, comprehension accuracy increases up to 85.5% and 96.3%. A closer inspection of the comprehension errors reveals a qualitative difference between the two kinds of semantic vectors. For simulated vectors, the majority of the errors (91.7%) involve lexemes, i.e., the recognized form has a different lexeme than the targeted form. For `fasttext` vectors, on the other hand, only about half of the errors are lexeme errors. The other half involves number errors, e.g., the singular form *minuta* 'minute' is recognized as its plural counterpart *minuti*. The reduced number of errors for held-out data in the model with simulated vectors suggest the orthogonality of the number features (singular, dual, plural) in the simulated semantic space is beneficial for generalization. However, the simulated vectors run the risk of oversimplifying the true complexity of plural semantics in Maltese (see, e.g., Shafaei-Bajestan et al., 2022, for English noun plurals).

For production, model performance with simulated and `fasttext` vectors is similar, at 68.3% and 64.4% respectively[4]. With simulated vectors, the correct form appears among the top 10 candidates for 72.8% of the held-

[3]The triphone models with simulated and `fasttext` vectors are provided in the supplementary material. The overall trend is similar, except that both comprehension and production accuracies are lower.

[4]For production of the held-out data, we lowered the threshold from 0.005 to 0.0005, and also for each word form, we allowed two cues to have support lower than the set threshold. This adjustment was motivated by the fact that some of the cues, due to their low frequency of occurrence in the training dataset, are not encountered often enough to be properly learned, and therefore require a more lenient criterion for acceptance as candidate cues for articulation.

out words; with `fasttext` vectors, this number increases to 90%. Table 4 displays the comprehension and production results for the held-out data:

	comprehension		production	
	simulated	fasttext	simulated	fasttext
top 1-candidate	77.8%	63.6%	68.3%	64.4%
top 10-candidates	85.5%	96.3%	72.8%	90%

Table 4: Model performance for comprehension (left) and production (right) for the held-out data. Rows indicate if only the predicted meaning (top 1-candidate) or the correct meaning among top 10-candidates was considered for the evaluation.

The majority of correctly produced forms belong to singular and sound plural forms, for both kinds of semantic vectors. Interestingly, in the case of broken plural forms we observe a different pattern: Among the 12 broken plural forms in the held-out dataset, the model using simulated vectors only produced one form correctly, while the model with `fasttext` vectors produced all 12 forms correctly. This may be due to the clustering of broken plurals in semantic space as gauged with fasttext (cf. Figure 2).

Further analyses on the production errors reveal that the type of errors that are made by simulated and `fasttext` models are also qualitatively different. Overall, we identified seven different error types for the production models that are shown in Table 5 below.

error type	simulated	`fasttext`	target	target lexeme	predicted	predicted lexeme
incorrect word	**84**	11	*ġar*	neighbor	*brejk*	brake
wrong affix	15	2	*satellita*	satellite	*satellitiku*	n.a.
phonetically close	10	**68**	*mera*	mirror	*mara*	woman
singular	5	29	*delegati*	delegates	*delegat*	delegate
plural	2	25	*minuta*	minute	*minuti*	minutes
alternative plural	2	0	*qlub*	cores	*qalbiet*	cores
missing diacritic	2	0	*rivalità*	rivalry	*rivalita*	n.a.
sum	120	135				

Table 5: Distribution of production errors from models with simulated and `fasttext` vectors along with examples for target word forms and their lexemes and predicted word forms and their lexemes (n.a. is given as lexeme for nonce-words).

In total, the two tested models produced 120 and 135 errors respectively. For most of the errors that the simulated model makes, 84 of 120 (70%) compared to 11 of 135 (8%) for the `fasttext` model, the predictions are far off from the targeted word forms. We labeled this category "incorrect word".

While some of these incorrect word forms are actual Maltese words, e.g. the model predicted *ġar* 'neighbor' for *brejk* 'brake', in some cases LDL produced new word forms, e.g. the pseudo-word *liku* for *ballu* 'dance (sg)'.

For 15 of 120 (12.5%) errors in the simulated model and 11 of 135 (1.5%) in the `fasttext` model, the models predicted a wrong affix. For example, the word *satellita* 'satellite' has the sound plural form, *satelliti*. While the target form was *satellita*, the models produced the non-existing form *satellitiku* using a wrong affix or additional phonological material, in this case -*ku*, for their predictions.

We labeled 10 of 120 (8.3%) errors in the simulated and 68 of 135 (50.4%) errors in the `fasttext` model as *phonetically close*. In these cases, LDL predicted a word form that is phonetically similar to the target word form, e.g. *mera* 'mirror' instead of *mara* 'woman'. For the `fasttext` model, these kind of errors concern the majority of all errors, thus highlighting the qualitative difference of the models' prediction again.

Other errors, 5 and 2 of 120 (4.2% and 1.6%) for the simulated model compared to 29 and 25 of 135 (21.5% and 18.5%) for the `fasttext` model involve mixing up singular and plural forms. For instance, the models predicted the singular *delegat* for *delegati* 'delegates'. Likewise, in another case the target word form was *minuta* 'minute' but the models predicted the plural form *minuti* 'minute' instead.

In a few cases, 2 of 120 (1.6%) errors for the simulated model (please note that this error did not occur at all in the `fasttext` model), LDL predicted an alternative plural form for a word that has multiple plural forms in our data set, e.g. *qalba* 'core (sg)' has three plural forms, one sound plural (*qalbiet*) and two broken plurals (*qlub* and *qliebi*). The testing data contained the broken plural form *qlub*, and the model predicted the sound plural *qalbiet* instead. This attraction to sound plurals is in line with the finding of a production study by Nieder, van de Vijver, et al. (2021a), in which native speakers tend to use frequent sound plural patterns for novel words.

The last minor group of errors, 2 of 120 (1.6%) of all errors in the simulated model (again, please note that this error did not occur at all in the `fasttext` model), concerns a missing diacritic. In two cases, the LDL prediction did not contain the diacritic of the target word form, e.g. *rivalita* for *rivalitá*.

4.3 Discussion

The explorations of Maltese noun inflection with LDL as computational engine for mappings between form and meaning clarified that model performance is excellent with the training data. For the held-out data, the model understands and produces unseen form with an accuracy around 70%, an accuracy that actually is surprisingly high for a noun system that is far from straightforwardly regular in many ways, and that can be expected to only be semi-productive. Compared to previous modeling results obtained within the framework of Word and Paradigm morphology (Nieder, Tomaschek, et al., 2021), accuracy is much higher than that of an Encoder-Decoder deep learning model, but lower than the exemplar-based model implemented with TiMBL. The TiMBL model, however, was given a much simpler task, namely, to predict classes of form changes, including classes bringing together many low-frequency patterns of change. In comparison to data from real speakers, the LDL model results on held-out data reflect the uncertainty of native speakers when it comes to infrequent words: Nieder, van de Vijver, et al. (2021a) asked participants to provide plural forms for given singulars, and observed that participants often were not able to provide the correct plural for existing infrequent singulars.

One modeling result is especially intriguing, namely, that to properly produce broken plurals for held-out data requires empirical, corpus-based vectors rather than simulated vectors. Conversely, simulated vectors outperform `fasttext` vectors when it comes to sound plurals. These observations suggest that there is a stronger isomorphism between the form space and the semantic space for the broken plurals.

We conclude that the theory of the Discriminative Lexicon, as a computational formalization of Word and Paradigm Morphology, provides a useful framework for predicting what forms are possible for listeners to understand, and what forms are possible for speakers to produce.

In what follows, we address the question of whether the way in which the discriminative lexicon model formalizes listening and speaking (admittedly at a high level of symbolic abstraction, especially when it comes to the representation of words' forms) can contribute to our understanding of human lexical processing. In the next section, we therefore examine whether measures derived from the model can contribute to enhancing statistical models fitted to response latencies in a primed lexical decision experiment with Maltese nouns that is reported in Nieder, van de Vijver, et al. (2021b).

5 Modeling Maltese priming reaction times

5.1 Maltese priming study

Nieder, van de Vijver, et al. (2021b) used a cross-modal priming paradigm with a lexical decision task to investigate the lexical storage and processing of Maltese sound and broken plurals.

In their study, they included 144 written singular targets from a Maltese noun list that appeared in one of two priming conditions: auditory primes were either 1) corresponding plural prime word forms, e.g. *klieb* - *KELB* ‘dogs - dog’, or, 2) phonologically and semantically unrelated control prime word forms that show the same plural suffix or pattern like the corresponding plural word, e.g. *bliet* - *KELB* ‘cities - dog’. They created two lists to prevent the same singular target appearing in both conditions for the same participant. In addition, Nieder, van de Vijver, et al. (2021b) included 144 nonce singular fillers created from existing Maltese singulars by changing the offset of the word forms (and thus keeping an initial phonological overlap with existing words). These nonce fillers were presented with the corresponding plural primes of the existing singulars that were used to create nonce words with, e.g. *klieb* - *KELT* ‘dogs - nonce filler’.

To investigate a possible frequency effect, Nieder, van de Vijver, et al. (2021b) reduced the Maltese plural variety substantially (see table 1 for an exhaustive list of the Maltese plural suffixes and patterns again) by including two frequent sound plural suffixes (*-i* and *-ijiet*), and two infrequent sound plurals (*-a* and *-at*), and two frequent broken plural templates (*CCVVCVC*, broken A and *CCVVC*, broken C) and two infrequent broken plural templates (*CCVjjVC*, broken D and *CCVVCV*, broken E). Their choice of including these plurals was motivated by the frequency results of a production study reported in Nieder, van de Vijver, et al. (2021a).

The results of the cross-modal priming study show no significant effect for plural type (*sound* vs. *broken*), but instead Nieder, van de Vijver, et al. (2021b) report that the reaction times of their participants were significantly influenced by the frequency of suffixes and patterns as well as by the word frequency of the singular targets. They conclude that Maltese sound and broken plurals are processed in the same way with pattern frequency being an important factor for lexical access.

5.2 Dataset

For exploring the usefulness of our computational model for understanding actual lexical processing, we re-analyzed the dataset from Nieder, van de Vijver, et al. (2021b). It contains 7885 observations (after removal of incorrect answers, practice trials and outliers) from fifty-nine participants.

In the following, when using the frequency of suffixes and templates as a variable for the model, we will use the terms “pattern frequency” and “patterns” to refer to both suffixes and templates. For the present study, we only used the reaction times for corresponding singular-plural pairs, omitting the control condition that was present in the experiment. Thus, to take the examples given above, we only included reaction times for *klieb* - *KELB* ‘dogs - dog’ but not for *bliet* - *KELB* ‘cities - dog’. This left us with 3995 observations. We then removed all words for which we did not have `fasttext` semantic vectors, resulting in a dataset with in all 2951 observations.

5.3 A baseline model

Extending the analyses of Nieder, van de Vijver, et al. (2021b), we predicted response times with `plural type` (TYPE), whether the plural form is sound or broken, and `pattern frequency` (PFREQ), whether the plural pattern is frequent or infrequent (cf. Table 1). In addition, we also included three lexical predictors pertinent to target words: `frequency` (FREQ), `neighborhood density` (ND)[5], and `word length` (LEN), measured in characters per word. In order to detect potential non-linear trends of the numeric predictors, we made use of the generalized additive mixed model (GAMM) provided by the `mgcv` package (v.1.8-36, Wood, 2017). The RTs (in seconds) were first inverse-transformed times -1 (so that small numbers still indicate fast RTs), and all the numeric predictors were log-transformed. In the model we allowed the two categorical factors to interact, and included by-subject random intercepts.

Word length did not contribute to improving model fit, and is therefore not considered in the analyses to follow. This is perhaps unsurprising, given that word length is highly correlated with neighborhood density ($r = 0.75$) and frequency ($r = -0.3$). A summary of the resulting model is presented in Table 6.

For this smaller dataset, re-analyzed with a GAMM instead of an LMM, and with ND as additional predictor, there is no evidence that RTs differ for broken and sound plurals, replicating the findings of Nieder, van de Vijver, et al. (2021b) for the original full dataset. The coefficient of pattern frequency indicates that plural primes with infrequent patterns induced longer response times compared to plural with frequent patterns. Figure 4 presents the partial effects of frequency and neighborhood density on RTs. The lines in the two plots center around zero because the intercepts and the adjustments of categorical predictors (part A in Table 6) are not included in the predictions: it is the effect of the predictor by itself that is shown. For frequency, the effect is nearly linear, but levels off for the highest-frequency words, a pattern often observed for word frequency in lexical decision tasks

[5]Neighborhood density is calculated on the basis of the vocabulary of `fasttext`, the size of which is about 120 thousand words, with punctuation and hyphenated words excluded.

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	-1.7891	0.0380	-47.1069	< 0.0001
$\text{TYPE}_{\texttt{sound}}$	0.0158	0.0178	0.8915	0.3727
$\text{PFREQ}_{\texttt{infrequent}}$	0.0448	0.0179	2.5050	0.0123
$\text{TYPE}_{\texttt{sound}}$:$\text{PFREQ}_{\texttt{infrequent}}$	-0.0198	0.0287	-0.6877	0.4917
B. smooth terms	edf	Ref.df	F-value	p-value
s(FREQ)	2.6296	3.1974	16.5970	< 0.0001
s(ND)	3.4840	3.8494	6.7397	< 0.0001
s(participant)	56.3240	58.0000	35.5194	< 0.0001

Table 6: Summary of a GAMM fitted to inverse-transformed RTs (-1/RT), with `plural type`, `pattern frequency`, `target frequency` and `neighborhood density` as fixed-effect predictors, and by-participant random intercepts.

(see, e.g., Baayen et al., 2006).[6] The effect of neighborhood density, on the other hand, is much more wiggly, and almost U-shaped. With the increase of neighborhood density, RTs first decrease and then increase. This U-shaped pattern suggests that participants responded faster for more probable values of ND as found in the center of the ND distribution, and responded more slowly for atypical values of ND, as found for atypically low and atypically high values of ND.

5.4 Predicting reaction times with LDL predictors

For predicting reaction times with measures based on discriminative learning, we opted for using the model with bi-syllables as cues, and `fasttext` word embeddings as semantic vectors. Bi-syllables were used as cues due to the better performance for these cues in training (see Table 3 again). We used `fasttext` vectors because, unlike simulated vectors, as we demonstrated above, they are remarkably sensitive to semantic differences between stems, number, plural types (broken vs. sound), and gender. Moreover, recall that the model using `fasttext` vectors did not only produce qualitatively different production errors (see table 5 again) but, contrary to the model using simulated vectors, managed to arrive at correct predictions for all broken plurals in the held-out data.

[6]Model fit can be further improved by including by-target random intercepts. However, due to very high concurvity, this model becomes uninterpretable: The covariates do not explain anything that is not already explained by the word-specific random intercepts. In this model, as well as in the model reported below, we therefore did not include by-target random intercepts.

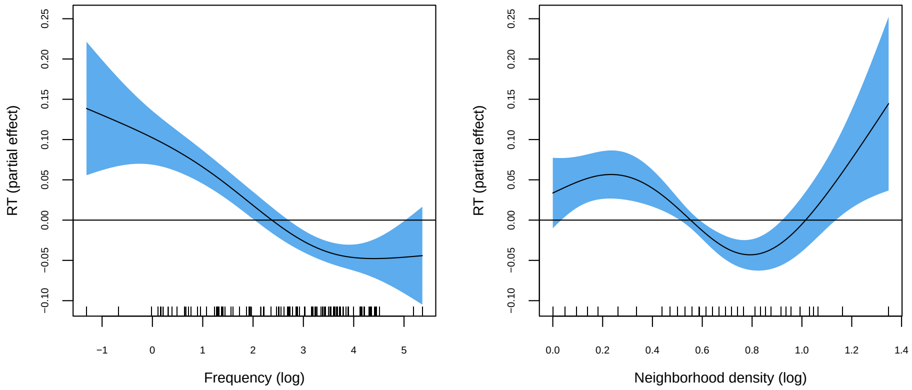

Figure 4: Effects of frequency and neighborhood density on RTs, indicated by the wiggly line. The blue area represent one standard error away from predictions. The rugs at the bottom indicate datapoint positions.

There are several potential measures that can be derived from an LDL model (see Chuang & Baayen, 2021, for an overview). We found two measures particularly useful for understanding Maltese unmasked primed lexical decision latencies, one measure quantifying how well primes' forms can be learned, and the other measure quantifying how closely the meaning of the prime plural already approximates the meaning of the target singular.

First consider the form measure, henceforth labeled `prime support`. The measure is defined as the sum of the positional semantic supports that the bi-syllable cues of a given plural prime word receive. By way of example, the word *trabi*, the plural form of *tarbija* (sg.f.) 'baby', contains three bi-syllable cues: `#.tra`, `tra.bi`, and `bi.#`, at positions 1, 2, and 3, respectively ("." denotes syllable boundaries). As described in Section 3.2.3, the `learn_paths` function in the **JudiLing** package calculates, for each cue position, the amount of support that a bi-syllable cue of the target word receives. That is, given the semantics of `trabi`, the positional support measure quantifies how certain the model is that `#.tra` should occur at position 1, `tra.bi` at position 2, and `bi.#` at position 3. For this example, the positional supports that the three bi-syllable cues receive are 0.25, 0.20, and 0.28, respectively. The `prime support` measure sums these three individual supports (i.e., $0.25 + 0.20 + 0.28 = 0.73$). The larger the `prime support` is, the more predictable a prime word's form is given its semantics, and the better its form is learned. This measure is motivated by two considerations. First, according to the motor theory of speech perception (Galantucci et al., 2006; Liberman & Mattingly, 1985), understanding the auditory prime necessarily involves internal production. Accordingly, the prime support measure, which captures the extent to which a word's temporally ordered triphones are supported by that word's semantics, is an integral part of the process of 'analysis by synthesis'. Empirical support for this measure is provided by Chuang, Kang, Luo, et al. (2021), who observed that the total positional support for Mandarin words was a co-determinant of their spoken word durations. Second, within the theory of the discriminative lexicon, internal comprehension is assumed to guide production (the previously introduced 'synthesis-by-analysis' approach, Baayen et al. (2018)). In this approach, comprehension and production are understood as more interlocked and interwoven than in classical models in which production and perception are allocated to encapsulated modules.

The second measure, henceforth labeled `pre-activation distance`, addresses the relation between plural prime words and their corresponding singular target words. It gauges the extent to which listening to a prime plural word semantically pre-activates (or "primes") the meaning of the target singular word. The `pre-activation distance` is defined as the Euclidean distance between the predicted semantic vector of a prime word and the gold standard semantic vector of its target word. A large value of this measure indicates that the predicted meaning of the plural prime word is far away in semantic space from the meaning of the singular target word. Conversely, a small pre-activation distance indicates that the prime word already closely approximates the meaning of the target word. This measure is inspired by a similar measure proposed in Baayen and Smolka (2020) on the basis of a naïve discrimination learning network, `prime-to-target pre-activation`, which calculates the extent to which a target word is already activated by the cues of the prime word. The current measure is modified to further take the semantics of prime and target words into account.

Similar to the baseline model, we fitted a GAMM to the inverse-transformed RTs with by-participant random intercepts, but this time with prime support and pre-activation distance as predictors. In addition, we asked a GAMM to predict the effects of both measures for sound and broken plurals separately, given that, as shown in Figure 5, in contrast to frequency and neighborhood density (top panel), a plural type difference emerges in the model and is naturally captured by the LDL measures (bottom panel), though the difference is more obviously pronounced for prime support than for pre-activation distance. The summary of the resulting model is presented in Table 7, and Figure 6 visualizes the partial effects of the two measures for sound and broken plurals.

For prime support (top panel), if we focus on where most datapoints are (indicated by rugs at the bottom of each figure), for both sound and broken plurals, the effect emerges as roughly inverse-U shaped: with increasing prime

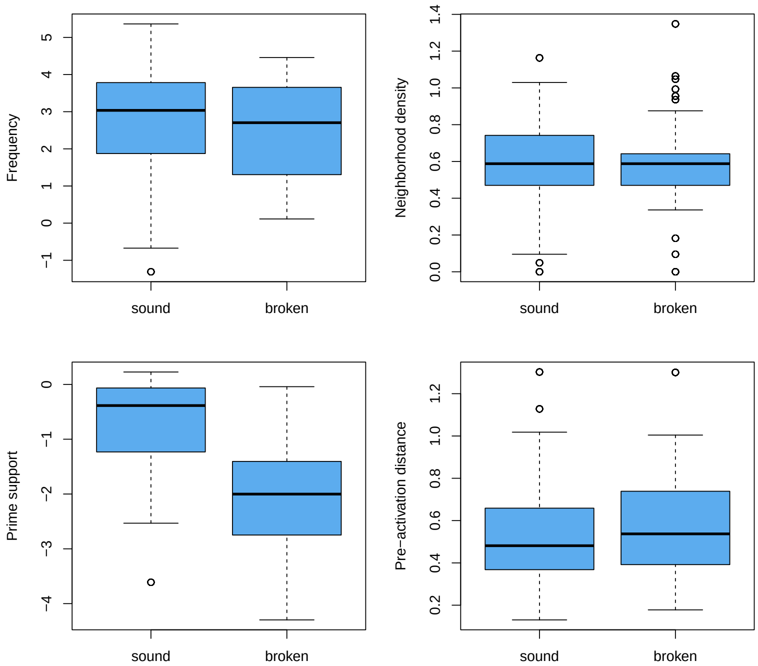

Figure 5: Boxplots of classical measures (frequency and neighborhood density, top panel) and LDL measures (prime support and pre-activation distance, bottom panel).

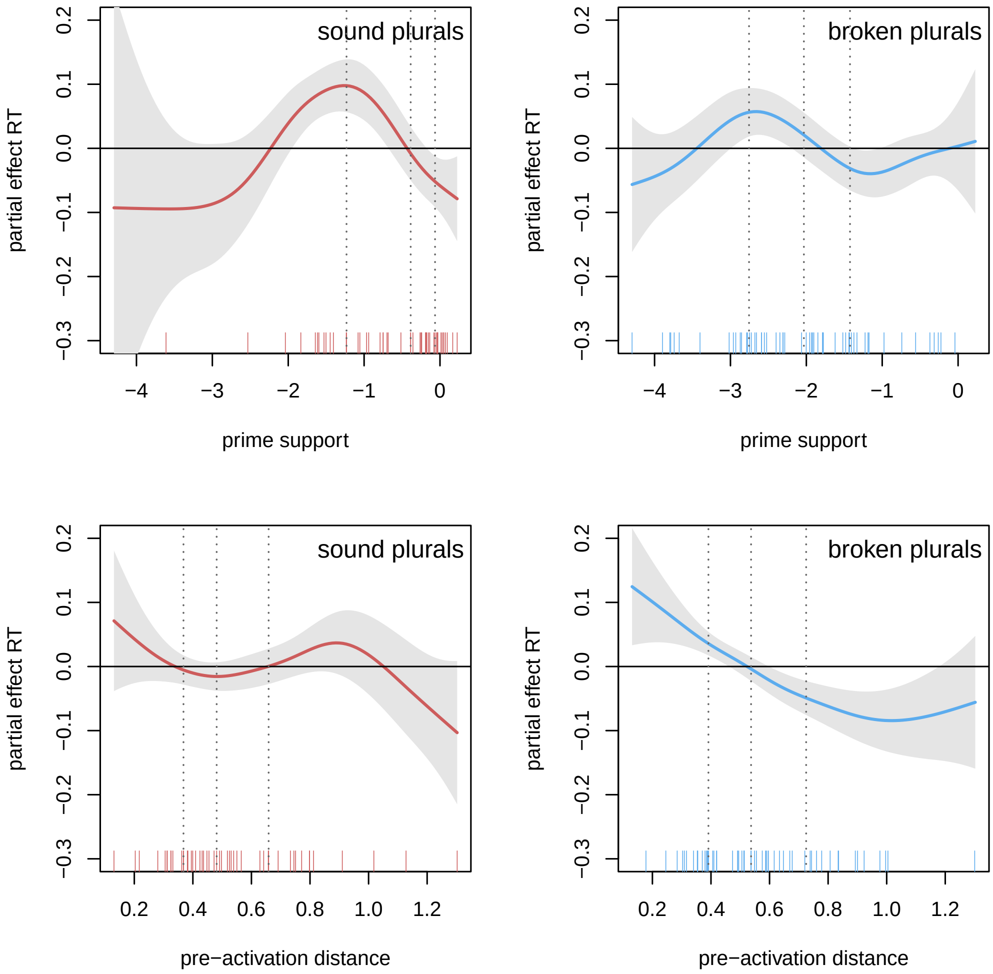

Figure 6: The effects of `prime support` (upper panel) and `pre-activation distance` (lower panel) on RTs for sound (left column) and broken (right column) plurals. The rugs at the bottom of each sub-figure indicate datapoint positions. The three dotted vertical lines in each sub-panel denote the first, second, and third quartiles.

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
(Intercept)	-1.7647	0.0382	-46.1578	< 0.0001
$\text{TYPE}_{\texttt{sound}}$	-0.0087	0.0205	-0.4236	0.6719
B. smooth terms	edf	Ref.df	F-value	p-value
s(primeSupport):$\text{TYPE}_{\texttt{sound}}$	3.8939	4.6880	10.5593	< 0.0001
s(primeSupport):$\text{TYPE}_{\texttt{broken}}$	4.1001	5.0081	4.1433	0.0009
s(preActDist):$\text{TYPE}_{\texttt{sound}}$	3.5923	4.4435	1.6615	0.1400
s(preActDist):$\text{TYPE}_{\texttt{broken}}$	2.6150	3.2798	6.8640	0.0001
s(participant)	56.3306	58.0000	35.6625	< 0.0001

Table 7: Summary of a GAMM fitted to inverse-transformed RTs, with the by-type smooths for `prime support` and `pre-activation distance` as fixed-effect predictors, and by-participant random intercepts.

support, RTs first increase and then decrease. Interestingly, the peaks of the inverse-U shape effects for both plural types coincide with their respective first quartile (25^{th} percentile). This suggests that for three quarters of the data, plurals that can be well predicted by semantics prime their singulars to a larger extent, resulting in shorter RTs. This pattern of results is in line with the effect of `prime-to-target pre-activation` as reported in Baayen and Smolka (2020). The trend, however, reverses for 25% of the plurals that are least learnable from their semantics. The inverse U-shaped curves suggest a trade-off between not knowing the prime's pronunciation, which makes it more like a pseudo-word, and knowing the prime's pronunciation well, which makes it more like a real word and thus enabling faster responses. How these two forces are balanced, and why the slowest responses are found at the first quartile, is unclear to us.

With respect to pre-activation distance, the effect is only seen for broken plurals. It is nearly linear with RTs becoming shorter as pre-activation distance increases. At first sight, the trend is puzzling, as one might have expected that if the prime fails to pre-activate the target, reaction times should be longer, but in reality, they are shorter. To make sense of this effect, we need to take a step back and have a critical look at the priming paradigm. Priming is often understood as involving facilitation of lexical access to the target. However, in general, compared to an identity baseline, primes typically give rise to longer instead of shorter response latencies. Primes are only facilitating when they are compared to an unrelated control baseline. In other words, unrelated primes are more disruptive than related primes, and related primes are more disruptive than identity primes.

The interpretation of primes as disrupting and interfering with normal lexical processing is supported by the experiments reported by Libben et al. (2018). Their study made use of primed visual lexical decision, with two-constituent compounds as target words and one of their constituents as primes. They observed longer reaction times for more frequent primes, in combination with the usual shorter reaction times for more frequent target compounds. In other words, their experiment indicates that the more frequent a prime is, the more it disrupts the processing of the target (see also Andrews, 1997; Forster & Hector, 2002, for interferences in priming and lexical retrieval).

With respect to the present experiment, a smaller pre-activation distance likewise bears witness to a similar disruptive effect of the prime. Since the plural and singular of a word are semantically highly similar in the first place, they are thus highly confusable and render deciding on the target's lexicality in general difficult. And as the prime and target are semantically more similar, the smaller the semantic distance of a prime to its target, the slower participants were able to make a lexicality decision, thus leading to longer RTs. Such a disruptive effect is more pronounced in broken plurals. The presence of a plural suffix in sound plurals (cf Table 1) possibly alleviates processing difficulties that arise when prime and target are very similar in meaning. Under close semantic proximity, sound plurals, thanks to the presence of a suffix, are easier to distinguish from their targets than broken plurals, which are more likely to be similar to simple words.

5.5 Discussion

How does the GAMM with LDL predictors compare to the baseline model with pattern frequency, target frequency and neighborhood density as predictors? To address this question, we compared Akaike's Information Criterion (AIC) for the two models. The AIC of the baseline model is 1900, and that of the LDL-based model is 1880. The corresponding evidence ratio is 22026, indicating that the LDL-based GAMM is 22026 times more likely than the baseline model to minimize the information loss.

We did not include target frequency as a predictor in the GAMM with prime support and pre-activation distance, for two reasons. First, within the framework of the discriminative lexicon, there are no word units with which frequency counts can be associated. Second, for modeling, we have made use of the multivariate multiple regression method for estimating weights, which represents the endstate of learning. At the endstate of learning, for which all token frequencies have increased to infinity, frequency effects are no longer present (see Heitmeier et al., 2021; Shafaei-Bajestan et al., 2021, for a detailed discussion).

Frequency of occurrence does come into play when incremental learning algorithms are used. For the present study, we did not explore incremental learning, for two reasons. First, for representing words' meanings, we would need incrementally updated semantic vectors. Unfortunately, incrementally updated `fasttext` vectors are not available for Maltese. Second, although incremental updating of the network is implemented in the **JudiLing** package for comprehension, it is not fully implemented for production. Developing a fully-fledged incremental version of the model is a target for further research. We do note, however, that when target frequency is added as predictor to the GAMM with LDL predictors, while prime support remain significant, pre-activation distance does not, and the effect size of target frequency reduces substantially. This is due to the high correlation between pre-activation distance and target frequency ($r = 0.62$), resulting in high concurvity and rendering the effects uninterpretable. Similarly, concurvity increases with neighborhood density added to the LDL model, as it is also highly correlated with pre-activation distance ($r = -0.63$).

It is noteworthy that in the baseline model with classical predictors, the type of prime was not supported as a predictor. In the model with LDL measures as predictors, an effect of the prime is detected, albeit not the originally anticipated effect of priming by plural type. In fact, according to this model, both a property of the prime (its 'pronouncability'), and the semantic relation of the prime to the target (gauged with `preActDist`), are the crucial predictors for participants' lexicality decision making.

6 General Discussion

We conclude this study with a discussion of the new insights that our results bring to morphological theory on the one hand, and the limitations of our approach on the other.

The semi-productivity of the Maltese plural poses a challenge for computational modeling. Any system, whether based on rules, analogy, or machine learning, needs to strike a balance between providing a good memory for the forms in use, and doing justice to the extent that the system is productive. We have shown that the Discriminative Lexicon (DL) model finds such a balance: it provides an accurate memory for both the comprehension and production of known words, and it also performs reasonably well when given the task to produce or understand novel, unseen forms. Given the semi-productivity of the Maltese plural, it is actually surprising how well prediction for unseen words works. This finding supports earlier descriptive studies that have called attention to substantial regularities in the Maltese plural system (Mayer et al., 2013; Nieder, van de Vijver, et al., 2021a; Schembri, 2012).

The theory of the DL currently does not include algorithms implementing decision making in experimental tasks such as lexical decision. Nevertheless, some headway can be made by incorporating measures derived from the theory as predictors in statistical models for experimental measures such as reaction times. Two such measures, one gauging how well we know a word's form, and the other assessing how closely the meaning of the prime approximates the meaning of the target, were found to improve the quality of a GAMM model fitted to the reaction times in a primed lexical decision task. The resulting model forced us to reconsider how to understand priming. Instead of understanding primes as facilitating lexical access to the target, primes may actually be disruptive. Among highly semantically relevant singular-plural word pairs, primes that are less similar in meaning to the target give rise to reduced interference. It should be kept in mind, however, that these results are tentative, based as they are on a post-hoc reanalysis, using exploratory data analysis, of the experiment reported earlier by Nieder, van de Vijver, et al. (2021b), and further replication studies will be essential for consolidating the present findings.

From this set of results, we conclude that the algorithm of linear discriminative learning, previously tested on Latin (Baayen et al., 2018), Estonian (Chuang et al., 2020), English (Chuang, Vollmer, et al., 2021), German (Heitmeier et al., 2021), Indonesian (Denistia & Baayen, 2021), Kinyarwanda (van de Vijver et al., 2021), and Korean (Chuang, Kang, Luo, et al., 2021), also provides a fruitful window on non-concatenative morphology.

The approach to the Maltese plural system that we have worked out in this study, which is a computational implementation of Word and Paradigm morphology (Blevins, 2016), differs from previous studies using computational modeling in that both production and comprehension are modeled. Instead of defining the task of morphological theory as providing a formal mechanism specifying what sound sequences are possible meaningful words, the DL framework explicitly addresses two challenges, first, to predict what possible forms are, given their meanings; and second, to predict what possible meanings are, given their forms. The present study is limited by the fact that the form representations that we made use of are based on abstract sublexical features such as letter or syllable n-grams, and it is currently an open question how the model will perform when, for instance, features derived from the acoustic signal are used (see Shafaei-Bajestan et al., 2021, for an exploration).

Our study also contributes to the theory of morphological productivity. Productivity is usually investigated for specific affixes. We have shown that we can assess the productivity of a whole system by inspecting how well the model's networks generalize to understanding and producing unseen forms. Several researchers have suggested that the productivity of rival affixes (e.g., -al, -ion, -ment) should be assessed jointly (Corbin, 1983; Wurzel, 1970; Zwanenburg, 1983). The present model for Maltese provides one way in which this suggestion can be implemented: many different suffixes for sound plurals, and many different templates for broken plurals,are all considered jointly.

Inspection of the semantics of Maltese singulars and plurals, using distributional semantics, clarified that the broken plurals, sound plurals, and singulars form partly overlapping but distinguishable clusters in semantic space. Furthermore, feminine and masculine nouns show some clustering in semantic space that is slightly different for singulars and plurals. These results show that the semantic vectors of inflected words have considerably more structure than expected in approaches in which plural inflection realizes a fixed morpho-syntactic feature. As the semantic vectors that can be simulated for inflected words with the **JudiLing** package implement fixed shifts for a given morpho-syntactic feature, it is clear that such vectors capture only part of the true complexity and richness of the semantics of inflected words. Simulated vectors construct a useful scaffolding for inflected words' semantics, sufficient to set up effective mappings between form and meaning, but insufficient for modeling the details of how form and meaning interact.

Since all models, including the one we presented in this paper, are idealizations, it is useful and necessary, we think, to reflect upon the differences between our model and native speakers of Maltese. The input to our model is a list of words and their semantics, conceptualized as embeddings. The model assumes that these forms and meanings are correct for any given speaker, but, of course, this is an idealization given that actual usage varies across speakers (Bybee, 2010; Sinclair, 1991b). Native speakers reported to us that they frequently hear other speakers use plurals that they had not heard before, but find understandable nevertheless (J. Nieder, personal communication, 2019).

Whereas native speakers learn continuously and incrementally, we have modeled the endstate of learning, of a learner with perfect memory and undivided attention to nouns alone. Obviously the existence of such a learner is a myth. It is possible to model incremental learning in LDL, but we do not have a sufficient amount of learning data of Maltese nouns to reliably model their learning. We leave this open for further research.

Our model represents a single (mythical) learner, but in reality there are individual differences between learners. Milin et al. (2017), for example, found evidence from skilled Russian readers that some readers accelerated as they progressed in a new text, whereas others slowed down. They connected this behavior to individual differences in the use of perceptual cues. Such individual differences in the use of cues would also affect acquisition of Maltese nouns. This could be modeled by learner-specific thresholds determining the number of candidate forms a speaker is willing to take into consideration (see also Chuang et al., 2020).

Keeping its limitations in mind, we contend that our model is useful as a quantitative tool for investigating high-level properties of human learning. Our model not only goes beyond predicting possible forms given another form, as is usual in computational models of morphophonology, but also provides model-based measures that predict human processing.

We conclude with noting that the LDL learning engine of the DL model strives for simplicity and interpretability. Formally, this engine carries out multivariate multiple linear regression on form and meaning. The assumption that mappings between form and meaning are linear undoubtedly involves substantial simplifications. Nevertheless, as illustrated in the present study, this simple approach already works surprisingly well, suggesting that the noun system of Maltese itself is also roughly 'linear'. Because the architecture of the network is fixed, and because there are very few hyperparameters (such as the threshold parameter that has to be set for production), the performance of the model is almost completely determined by the representations selected by the researcher for representing form and meaning, and the data. This, we think, makes the model especially useful as a tool for linguistic analysis.

Data Availability Statement

The data that support the findings of this study are openly available at (anonymous view-only link):
https://osf.io/rxsbu/?view_only=73d7c8d1fb854592a293d891383d7e7a.

References

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

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review*, (4), 439–61. https://doi.org/https://doi.org/10.3758/BF03214334

Aquilina, J. (1987). *Maltese-English dictionary*. Midsea Books.

Baayen, R. H. (2011). Corpus linguistics and naive discriminative learning. *Brazilian Journal of Applied Linguistics*, *11*, 295–328.

Baayen, R. H., Chuang, Y.-Y., & Blevins, J. P. (2018). Inflectional morphology with linear mappings. *The Mental Lexicon*, *13*(2), 230–268.

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., Feldman, L., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language*, *53*, 496–512.

Baayen, R. H., & Moscoso del Prado Martín, F. (2005). Semantic density and past-tense formation in three Germanic languages. *Language*, *81*, 666–698.

Baayen, R. H., & Smolka, E. (2020). Modeling morphological priming in German with naive discriminative learning. *Frontiers in Communication*, *5*, 17.

Baroni, M., Dinu, G., & Kruszewski, G. (2014). Don't count, predict! a systematic comparison of context-counting vs. context-predicting semantic vectors. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 238–247.

Belth, C., Payne, S., Beser, D., Kodner, J., & Yang, C. (2021). The greedy and recursive search for morphological productivity. *arXiv*. https://arxiv.org/abs/2105.05790

Berko, J. (1958). The child's learning of english morphology. *WORD*, *14*(2), 150–177. https://doi.org/10.1080/00437956.1958.11659661

Bird, H., Lambon Ralph, M., Seidenberg, M., McClelland, J., & Patterson, K. (2003). Deficits in phonology and past-tense morphology: What's the connection? *Journal of Memory and Language*, *48*, 502–526.

Blevins, J. P. (2016). The minimal sign. In A. Hippisley & G. Stump (Eds.), *The Cambridge Handbook of Morphology* (pp. 50–69). Cambridge University Press. https://doi.org/10.1017/9781139814720.003

Boersma, P., & Hayes, B. (2001). Empirical tests of the gradual learning algorithm. *Linguistic Inquiry*, *32*, 45–86.

Boersma, P. (1998). *Functional Phonology* (Doctoral dissertation). University of Amsterdam.

Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching word vectors with subword information. *arXiv*. https://arxiv.org/abs/1607.04606

Bonami, O., & Stump, G. T. (2016). Paradigm function morphology. In A. Hippisley & G. T. Stump (Eds.), *Cambridge Handbook of Morphology* (pp. 449–481). Cambridge University Press.

Booij, G. E. (1996). Inherent versus contextual inflection and the split morphology hypothesis. In G. E. Booij & J. v. Marle (Eds.), *Yearbook of morphology 1995* (pp. 1–16). Kluwer Academic Publishers.

Borg, A., & Azzopardi-Alexander, M. (1997). *Maltese*. Routledge.

Bybee, J. L. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, *10*(5), 425–455. https://doi.org/10.1080/01690969508407111

Bybee, J. L. (2010). *Language, usage and cognition*. Cambridge University Press.

Camilleri, J. J. (2013). *A Computational Grammar and Lexicon for Maltese* (Master Thesis). University of Gothenburg. Gothenburg, Sweden.

Chuang, Y.-Y., & Baayen, R. H. (2021). Discriminative learning and the lexicon: NDL and LDL. *Oxford Research Encyclopedia of Linguistics* (in press). Oxford University Press.

Chuang, Y.-Y., Kang, M., Luo, X. F., & Baayen, R. H. (2021). Vector space morphology with linear discriminative learning. In D. Crepaldi (Ed.), *Linguistic morphology in the mind and brain*. Routledge.

Chuang, Y.-Y., Kang, M., Luo, X., & Baayen, R. H. (2021). Vector space morphology with linear discriminative learning. *arXiv*. https://arxiv.org/abs/2107.03950

Chuang, Y.-Y., Lõo, K., Blevins, J. P., & Baayen, R. H. (2020). Estonian Case Inflection Made Simple. A Case Study in Word and Paradigm Morphology with Linear Discriminative Learning. In Körtvélyessy, Livia and Štekauer, Pavol (Ed.), *Complex Words: Advances in Morphology* (pp. 119–141). Cambridge University Press.

Chuang, Y.-Y., Vollmer, M.-L., Shafaei-Bajestan, E., Gahl, S., Hendrix, P., & Baayen, R. H. (2021). The processing of pseudoword form and meaning in production and comprehension: A computational modeling ap-

proach using linear discriminative learning. *Behavior Research Methods*, *53*(3), 945–976. https://doi.org/10.3758/s13428-020-01356-w
Corbin, D. (1983). Le monde étrange des dictionnaires (4): La créativité lexicale, le lexicographe et le linguiste. *Lexique*, *2*, 43–68.
Creemers, A., Davies, A. G., Wilder, R. J., Tamminga, M., & Embick, D. (2020). Opacity, transparency, and morphological priming: A study of prefixed verbs in dutch. *Journal of Memory and Language*, *110*.
Daelemans, W., & Van den Bosch, A. (2005). *Memory-based language processing*. Cambridge University Press.
Daelemans, W., Zavrel, J., van der Sloot, K., & van den Bosch, A. (2001). TiMBL: Tilburg Memory-Based Learner-version 4.1-Reference Guide.
Daelemans, W., Zavrel, J., Van Der Sloot, K., & Van den Bosch, A. (2004). *Timbl: Tilburg memory-based learner* (tech. rep.). Tilburg University.
Dawdy-Hesterberg, L. G., & Pierrehumbert, J. B. (2014). Learnability and generalisation of Arabic broken plural nouns. *Language, Cognition and Neuroscience*, *29*(10), 1268–1282. https://doi.org/10.1080/23273798.2014.899377
Dell, G. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological review*, *93*(3), 283–321. https://doi.org/10.1037/0033-295X.93.3.283
Denistia, K., & Baayen, R. H. (2021). The morphology of indonesian: Data and quantitative modeling. In C. Shei (Ed.), *The Routledge Handbook of Asian Linguistics*. Routledge.
Drake, S. (2018). The form and productivity of the Maltese morphological diminutive. *Morphology*, *28*(3), 297–323. https://doi.org/https://doi.org/10.1007/s11525-018-9328-0
Ernestus, M., & Baayen, R. H. (2003). Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language*, *79*, 5–38.
Farrugia, A., & Rosner, M. (2008). Maltimorph, a computational analysis of the Maltese broken plural. *Proceedings of Workshop on ICT 2008*.
Forster, K. I., & Hector, J. (2002). Cascaded versus noncascaded models of lexical and semantic processing: The turple effect. *Memory and Cognition*, *30*(7), 1106–1117. https://doi.org/https://doi.org/10/dgzskq
Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic bulletin & review*, *13*(3), 361–377.
Gaskell, M. G., & Marslen-Wilson, W. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and Cognitive Processes*, *12*, 613–656. https://doi.org/10.1080/016909697386646

Gatt, A., & Čéplö, S. (2013). Digital corpora and other electronic resources for Maltese. *Proceedings of the International Conference on Corpus Linguistics*.

Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, *111*, 662–720.

Heitmeier, M., & Baayen, R. H. (2021). Simulating phonological and semantic impairment of English tense inflection with Linear Discriminative Learning. *The Mental Lexicon*, *15*, 385–421.

Heitmeier, M., Chuang, Y.-Y., & Baayen, R. H. (2021). Modeling morphology with linear discriminative learning: Considerations and design choices.

Hoberman, R. D. (2007). Maltese morphology. In A. S. Kaye (Ed.), *Morphologies of asia and africa* (pp. 257–281). Eisenbrauns.

Ivens, S. H., & Koslin, B. L. (1991). *Demands for reading literacy require new accountability methods*. Touchstone Applied Science Associates.

Joulin, A., Grave, E., Bojanowski, P., Douze, M., Jégou, H., & Mikolov, T. (2016). Fasttext.zip: Compressing text classification models. *arXiv*. https://arxiv.org/abs/1612.03651

Joulin, A., Grave, E., Bojanowski, P., & Mikolov, T. (2016). Bag of tricks for efficient text classification. *arXiv*. https://arxiv.org/abs/1607.01759

Kastner, I. (2019). Templatic morphology as an emergent property. *Natural Language & Linguistic Theory*, *37*(2), 571–619. https://doi.org/10.1007/s11049-018-9419-y

Klafehn, T. (2013). Myth of the wug test: Japanese speakers can't pass it and English-speaking children can't pass it either. In C. Cathcart, I.-H. Chen, G. Finley, S. Kang, C. S. Sandy, & E. Stickles (Eds.), *Proceedings of the 37th annual meeting of the berkely linguistics society* (pp. 170–184).

Kunter, G. (2017). Coquery: A free corpus tool. version 0.10.0.. https://www.coquery.org

Landauer, T., & Dumais, S. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction and representation of knowledge. *Psychological Review*, *104*(2), 211–240.

Levelt, W., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1–38.

Libben, G., Goral, M., & Baayen, R. H. (2018). What does constituent priming mean in the investigation of compound processing? *The Mental Lexicon*, *13*(2), 269–284.

Liberman, A., & Mattingly, I. (1985). The motor theory of speech perception revised. *Cognition*, *21*, 1–36.

Luo, X., Chuang, Y.-Y., & Baayen, R. H. (2021). Judiling: An implementation in Julia of Linear Discriminative Learning algorithms for language modeling. https://megamindhenry.github.io/JudiLing.jl/stable/

Matthews, P. H., & Matthews, P. (1972). *Inflectional morphology: A theoretical study based on aspects of Latin verb conjugation* (Vol. 6). CUP Archive.

Mayer, T., Spagnol, M., & Schönhuber, F. (2013). Fixing the broken plural in Maltese. *Perspectives on Maltese Linguistics* (pp. 129–158). Akademie Verlag.

McCarthy, J. J. (1981). A Prosodic Theory of Non-concatenative Morphology. *Linguistic Inquiry*, *12*, 373–418.

McCarthy, J. J., & Prince, A. S. (1990). Foot and word in prosodic morphology: The arabic broken plural. *Nat Lang Linguist Theory*, *8*(2), 209–283. https://doi.org/10.1007/BF00208524

McCarthy, J. J., & Prince, A. S. (1996). Prosodic morphology. *Linguistics Department Faculty Publication Series*.

McCoy, R. T., Frank, R., & Linzen, T. (2020). Does Syntax Need to Grow on Trees? Sources of Hierarchical Inductive Bias in Sequence-to-Sequence Networks. *Transactions of the Association for Computational Linguistics*, *8*, 125–140. https://doi.org/10.1162/tacl_a_00304

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, 3111–3119.

Milin, P., Divjak, D., & Baayen, R. H. (2017). A learning perspective on individual differences in skilled reading: Exploring and exploiting orthographic and semantic discrimination cues. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(11), 1730.

Mirković, J., MacDonald, M. C., & Seidenberg, M. S. (2005). Where does gender come from? evidence from a complex inflectional system. *Language and cognitive processes*, *20*, 139–167.

Mitchell, J., & Lapata, M. (2008). Vector-based models of semantic composition. *ACL*, 236–244.

Nieder, J., Tomaschek, F., Cohrs, E., & van de Vijver, R. (2021). Modelling Maltese noun plural classes without morphemes. *Language, Cognition and Neuroscience*. https://doi.org/10.1080/23273798.2021.1977835

Nieder, J., van de Vijver, R., & Mitterer, H. (2021a). Knowledge of Maltese singular–plural mappings. *Morphology*, *31*, 147–170.

Nieder, J., van de Vijver, R., & Mitterer, H. (2021b). Priming Maltese plurals: Representation of sound and broken plurals in the mental lexicon. *The Mental Lexicon*, *16*(1), 69–97.

Nosofsky, R. (1986). Attention, similarity and the identification-categorization relationship. *Experimental Psychology: General*, *115*(1), 39–57.

Pater, J. (2004). Bridging the gap between receptive and productive development with minimally violable constraints. In R. Kager, J. Pater, & W. Zonneveld (Eds.), *Constraints in Phonological Acquisition* (pp. 219–244). Cambridge University Press.

Pennington, J., Socher, R., & Manning, C. D. (2014). Glove: Global vectors for word representation. *Empirical Methods in Natural Language Processing (EMNLP)*, 1532–1543. http://www.aclweb.org/anthology/D14-1162

Prince, A., & Smolensky, P. (2004). *Optimality Theory: Constraint Interaction in Generative Grammar*. Blackwell.

Rumelhart, D. E., & McClelland, J. L. (1986). On Learning the Past Tenses of English Verbs. In J. L. McClelland & D. E. Rumelhart (Eds.), *Parallel Distributed Processing. Explorations in the Microstructure of Cognition. Vol. 2: Psychological and Biological Models* (pp. 216–271). The MIT Press.

Saito, M., Tomaschek, F., & Baayen, R. (2021). Relative functional load determines co-articulatory movements of the tonguetip. In M. Tiede, D. H. Whalen, & V. Gracco (Eds.), *Proceedings of the 12th International Seminar on Speech Production (ISSP 2020)*. Haskins Press.

Schembri, T. (2012). *The broken plural in Maltese—A description.* Brockmeyer.

Shafaei-Bajestan, E., Moradipour-Tari, M., Uhrig, P., & Baayen, R. H. (2021). LDL-AURIS: A computational model, grounded in error-driven learning, for the comprehension of single spoken words. *Language, Cognition and Neuroscience.* https://doi.org/https://www.tandfonline.com/doi/full/10.1080/23273798.2021.1954207

Shafaei-Bajestan, E., Moradipour-Tari, M., Uhrig, P., & Baayen, R. H. (2022). Semantic properties of english nominal pluralization: Insights from word embeddings. *Manuscript, University of Tübingen.*

Sinclair, J. (1991a). *Corpus, concordance, collocation.* Oxford University Press.

Sinclair, J. (1991b). *Corpus, concordance, collocation.* Oxford University Press.

Skousen, R. (1989). *Analogical modeling of language.* Kluwer.

Smolensky, P. (1996). On the comprehension/production dilemma in child language. *Linguistic inquiry*, *27*(4), 720–731.

Stump, G. T. (2001). *Inflectional morphology: A theory of paradigm structure* (Vol. 93). Cambridge University Press.

Tomaschek, F., Plag, I., Ernestus, M., & Baayen, R. H. (2021). 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. https://doi.org/10.1017/S0022226719000203

Ussishkin, A. (2005). A fixed prosodic theory of nonconcatenative templatic morphology. *Natural Language & Linguistic Theory*, *23*(1), 169–218. http://www.jstor.org/stable/4048019

van de Vijver, R., & Baer-Henney, D. (2014). Developing biases. *Frontiers in Language Sciences*, *5*(634). https://doi.org/10.3389/fpsyg.2014.00634

van de Vijver, R., & Uwambayinema, E. (accepted). A word-based account of comprehension and production of Kinyarwanda nouns in the Discriminative Lexicon. *Linguistics Vanguard*.

van de Vijver, R., Uwambayinema, E., & Chuang, Y.-Y. (2021). Comprehension and production of Kinyarwanda verbs in the Discriminative Lexicon. *PsyArXiv*. https://doi.org/10.31234/osf.io/h63zy

Widrow, B., & Hoff, M. E. (1960). Adaptive switching circuits.

Wood, S. N. (2017). *Generalized Additive Models*. Chapman & Hall/CRC.

Wurzel, W. U. (1970). *Studien zur deutschen Lautstruktur*. Akademie-Verlag.

Yang, W., Lu, W., & Zheng, V. (2017). A simple regularization-based algorithm for learning cross-domain word embeddings. *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. https://doi.org/10.18653/v1/d17-1312

Zamuner, T., Kerkhoff, A., & Fikkert, P. (2006). *Children's knowledge of how phonotactics and morphology interact* [Ms. Radboud University Nijmegen, University of Britsh Columbia and Utrecht University].

Zamuner, T., Kerkhoff, A., & Fikkert, P. (2011). Phonotactics and morphophonology in early child language: Evidence from Dutch. *Applied Psycholinguistics*, *33*(3), 481–499. https://doi.org/10.1017/S0142716411000440

Zwanenburg, W. (1983). *Productivité Morphologique et Emprunt. Études des dérivés déverbaux savants en français moderne*. Benjamins.