The Multiple Inflection Generator: A generalized connectionist model for cross-linguistic morphological development.

Themelis N. Karaminis\textsuperscript{1,2}
Michael S. C. Thomas\textsuperscript{1}

\textsuperscript{1}Developmental Neurocognition Lab, Birkbeck, University of London
\textsuperscript{2}Centre for Research in Autism and Education, Institute of Education, University of London

Contact author:

Dr Themelis Karaminis
Centre for Research in Autism and Education (CRAE)
Department of Psychology and Human Development
Institute of Education, University of London
Room 402, 55-59 Gordon Square
London, WC1H 0NU
E-mail: t.karaminis@ioe.ac.uk
Telephone: +44 (0) 20 79115301
Abstract

The next challenge for connectionist models of the acquisition of inflectional morphology (IM) is to increase their generality: across inflectional paradigms, across grammatical classes, and ultimately across languages. We present a new model of IM that draws together elements of several existing connectionist models and which acquires multiple inflectional paradigms across three grammatical classes. Importantly, the same model was required to learn artificial languages capturing the key properties of two different types of languages: English, as an example of a morphologically poor language, and Modern Greek, as an example of a morphological rich language. The model was evaluated against its ability to simulate 11 empirical phenomena of IM acquisition in English, and 7 empirical phenomena of IM acquisition in Modern Greek. The model succeeded in capturing 16 of 18 in total. The emergent functional structure of the IM system depended on the statistical properties of the language to which the system was exposed. The model was most consistent with a multiple cues perspective on language acquisition. However, its account of IM acquisition also had similarities to the dual route (Pinker, 1991, 1994, 1999) and optional infinitive (Wexler, 1994, 1999) theories.

Keywords: cross-linguistic language acquisition, inflectional morphology, computational modeling, parallel distributed processing
Computational models have been useful in advancing our understanding of the mechanisms that underlie language development. In the domain of Inflectional Morphology (IM), for example, an extensive literature of artificial neural network or connectionist models offers mechanistic explanations for the emergence of a wide range of empirical phenomena, such as accuracy rates and error patterns in regular and irregular inflection, type and token frequency effects, and preferences for the inflection of novel items (e.g., Forrester & Plunkett, 1994; Hoeffner, 1992; Hoeffner & McClelland, 1993; Indefrey & Goebel, 1993; Joanisse, 2004; Joanisse & Seidenberg, 1999; MacWhinney & Leinbach, 1991; Mirković, Seidenberg, & Joanisse, 2011; Nakisa & Hahn, 1996; Plunkett & Juola, 1999; Plunkett & Marchman, 1991, 1993, 1996; Plunkett & Nakisa, 1997; Ruh & Westermann, 2008, 2009; Rumelhart & McClelland, 1986; Thomas, 2005; Thomas, Forrester, & Ronald, 2013; Thomas & Karmiloff-Smith, 2003; Thomas & Knowland, in press; Westermann & Ruh, 2012; Woollams, Joanisse, & Patterson, 2009). Such explanations typically rely upon the study of interactions between a learning system incorporating general properties of computations in the brain (i.e., a Parallel Distributed Processing/PDP architecture; Rumelhart, McClelland, & the PDP Research Group, 1986) and statistical properties of the linguistic environments (training sets) to which it has been exposed.

Despite their success, connectionist models of the acquisition of IM have on occasions been insufficiently general in three ways. First, models have targeted individual subdomains of IM rather than fully-fledged morphological systems. For example, many models focused on the past tense, under the assumption that this subdomain taps the main cognitive processes involved in the acquisition and use of morphological knowledge (e.g., Hare & Elman, 1995; Hare, Elman, & Daugherty,
1995; Indefrey & Goebel, 1993; Joanisse, 2004; Joanisse & Seidenberg, 1999;
Westermann, 2008, 2009; Rumelhart & McClelland, 1986; Thomas, 2005; Thomas &
Karmiloff-Smith, 2003; Thomas & Knowland, in press; Thomas et al., 2013;
Westermann & Ruh, 2012; Woollams et al., 2009). Second, connectionist models of
IM have focused on individual languages, and especially English (e.g., Hoeffner,
MacWhinney & Leinbach, 1991; Plunkett & Juola, 1999; Plunkett & Marchman,
1991, 1993, 1996; Rumelhart & McClelland; 1986; Thomas, 2005; Thomas &
Karmiloff-Smith, 2003; Thomas & Knowland, in press; Thomas, Forrester, & Ronald,
2013; Westermann & Ruh, 2012; Woollams et al., 2009). Third, most models have
addressed only typical development (e.g., Forrester & Plunkett, 1994; Hoeffner, 1992;
Indefrey & Goebel, 1993; MacWhinney & Leinbach, 1991; Mirković et al., 2011;
Nakisa & Hahn, 1996; Plunkett & Juola, 1999; Plunkett & Marchman, 1991, 1993,
1996; Plunkett & Nakisa, 1997; Ruh & Westermann, 2008, 2009; Rumelhart &
McClelland, 1986; Westermann & Ruh, 2012; Woollams et al., 2009) with few
models including deficits in developmental or acquired disorders (e.g., Hoeffner &
McClelland, 1993; Joanisse, 2004; Joanisse & Seidenberg, 1999; Thomas, 2005;
Thomas & Karmiloff-Smith, 2003, Thomas & Knowland, in press; Thomas et al.,
2013).

These limitations have implications for the plausibility of connectionist
accounts of morphological development. With regard to the focus of connectionist
models on individual subdomains of IM, one could argue that the presence of a
specific cognitive system dedicated to the processing of a particular inflection – e.g.,
past tense and not, say, progressive – is highly unlikely (cf. Plunkett & Juola, 1999,
also evidence from neuroimaging: Tyler, Bright, Fletcher, & Stamatakis, 2004; Yokoyama et al., 2006). With regard to the focus of models on individual languages, models should have no language-specific structures and should work across language typologies (cf. Hutzler, Ziegler, Perry, Wimmer, & Zorzi, 2004; Seidenberg, 2011 on reading development models). English, for example, has a simple morphological system, characterised by predominant regularity (e.g., Thomas & Karmiloff-Smith, 2003). This is not the case in other languages, such as Arabic (Forrester & Plunkett, 1994; Plunkett & Nakisa, 1997), French (Prevost, 2009), German (Nakisa & Hahn, 1996), Icelandic (Ragnarsdottir, Simonsen, & Plunkett, 1999), Modern Greek (Stephany, 1997), or Serbian (Mirković et al., 2011). The language generality of a model’s architecture cannot be tested unless it is applied to acquiring the IM of another language. Finally, regarding the focus of neural network models on typical language development, models accounting for the acquisition of IM based on specific assumptions on the nature of the underlying neurocomputational mechanisms should also show how potential impairments of these mechanisms may (or not) capture atypical linguistic profiles (e.g., Thomas & Karmiloff-Smith, 2003; Thomas & Knowland, in press; Thomas et al., 2013; see also Thomas, 2005).

When considered together, these issues pose a challenge for connectionist approaches to morphological development. This is to propose a neural network model of IM that is simultaneously general across three dimensions: 1) it implements a fully-fledged inflectional system, rather than a piece-meal model of a particular inflection type; 2) it applies to developmental phenomena in multiple languages; and 3) it addresses both typical and atypical development.

In this paper, we present a connectionist model for the acquisition of IM which represents a first step towards addressing this challenge and targets the first two
dimensions of generality, i.e., dimensions 1) and 2). Elsewhere, we show how the model is also general across the third dimension of typical and atypical development (Karaminis, 2012; Karaminis & Thomas, in preparation). The Multiple Inflection Generator (MIG) implements a scaled-up inflectional system comprising three grammatical classes (nouns, adjectives, and verbs) and multiple inflections within a grammatical class (e.g., English verbs: base forms, past tense, progressive, third person singular; English nouns: base forms, plural, genitive). At the same time, the MIG has a cross-linguistic dimension. It addresses behavioural phenomena in the acquisition of IM in two languages with different degrees of morphological richness, namely English and Modern Greek.

The MIG was inspired by previous connectionist models that addressed the acquisition of multiple inflections either within or across grammatical classes (multiple verb inflections; Hoeffner, 1992; Hoeffner & McClelland, 1993; MacWhinney & Leinbach, 1991; multiple noun inflections; Mirković et al., 2011; inflections of multiple grammatical classes: Plunkett & Juola, 1999). Another source of inspiration was models studying the acquisition of IM in non-English languages (Forrester & Plunkett, 1994; Hare & Elman, 1995; Hare et al., 1995; Indefrey & Goebel, 1993; Mirković et al., 2011; Nakisa & Hahn, 1996; Plunkett & Nakisa, 1997; Ruh & Westermann, 2008, 2009). Connectionist models of these two categories demonstrated the potential of the PDP framework to be general across dimensions 1) and 2) individually, that is, to simulate either phenomena in the acquisition of multiple inflections or phenomena in the acquisition of individual subdomains across languages. It remained to be shown that these two dimensions of generality could be combined in a single model. Such a computational model should be robust to interactions arising from the acquisition of multiple inflections of multiple
grammatical classes and robust to language typology, as a factor that determines how these interactions are manifested within a given system of IM. The new model should also have the power to unify an extensive literature of empirical effects in the acquisition of different inflections from different grammatical classes in English and, in this case, Modern Greek, under the general principles of the PDP framework. Importantly, it should be able to do so under a common set of theoretical assumptions and simplifications for the component of the language system supporting the acquisition and use of IM.

A key step in our research design was the development of two training sets representing the linguistic environment of the child acquiring English or Modern Greek as a first language (correspondingly). These reflected key characteristics of the system of IM in English and Modern Greek, as well as key cross-linguistic differences with respect to IM and phonology. The two training sets were used to train the same neural network architecture, with minor modifications only to accommodate cross-linguistic differences in phonology. English was modelled as a language making wide use of morphologically unmarked forms and employing a simple morphological system characterised by predominant regularity (cf. Thomas & Karmiloff-Smith, 2003). Modern Greek, on the other hand, was modelled as a language featuring obligatory morphological marking for nouns, adjectives, and verbs, and a rich system of IM that included conjugational classes (Stephany, 1997). An important part of the model was a frequency structure that reflected frequencies of grammatical classes, inflection types, regular and irregular paradigms, and conjugational classes within each language (type frequencies). This structure was largely based on measurements of text corpora (English: Francis & Kučera, 1982; Modern Greek: Hatzigeorgiu et al.,
2000), and was combined with a simplified two-level frequency scheme for individual exemplars (token frequency; high vs. low).

Our broader theoretical position is that the acquisition of IM involves learning to integrate multiple types of information (‘cues’: stem phonology, lexical semantics, grammatical class, and target inflection information) so as to produce appropriately inflected phonological word forms, in accordance with the grammatical context. We show that a simple feed-forward architecture receiving multiple cues as input and trained to produce phonological forms corresponding to appropriately inflected words in the output layer is able to learn training sets representing fully-blow morphological systems, either similar to English or to Modern Greek. We also show that this multiple-cue architecture acquires English and Modern Greek IM in a psycholinguistically plausible manner. We analyse results from simulations with the MIG to delineate how a large body of empirical effects in the acquisition of English and Modern Greek IM emerges through interactions between general properties of a PDP learning system (e.g., similarity-based processing of activation patterns) and statistical characteristics of the corresponding training sets, such as frequencies of different inflections and individual exemplars, the level of complexity of different inflections (e.g., progressive simpler than past tense), and similarities and differences between different types of mappings.

Methods and advances in connectionist models of IM

Table 1 summarises the main features of representative neural network models of IM. Rows in this table refer to individual models, starting from the Rumelhart and McClelland (1986) model for the English past tense and moving to more recent ones, such as the Mirković et al. (2011) model for the inflections of Serbian nouns. The
second column of Table 1 describes the scope of the models, i.e., the natural language, the target domain (e.g., plural) and the cohort (e.g., typically developing children) to which the target empirical phenomena referred. The third column includes information on the training set (number of mappings) and how the linguistic environment was modelled, e.g., using training sets corresponding to real vs. artificial words, or adopting incremental vs. non-incremental training regimes (see below in this section). Finally, the fourth column refers to the main types of information included in the neural network architecture of a model, as well as the types of representations (localist/distributed), the types of connections (uni-/bidirectional), and the type of network (e.g., feed-forward/ attractor/ constructivist).

We will refer to Table 1 and to features of individual models in order to outline major assumptions of the connectionist approach with respect to to five issues: 1) the representation of phonological information; 2) the representation of lexical-semantic information; 2) the encoding of the linguistic environment; 4) the inclusion of multiple inflection types; and 5) the modeling of non-English IM.

Distributed phonological representations: From Wickelfeatures to sequences of phonemes

As shown in Table 1, the Rumelhart and McClelland (1986) model differed from all other models of IM in its use of Wickelfeature phonological representations (abbreviated as P^WF in the last column). The implausibility of these representations was a major point of criticism against Rumelhart and McClelland (1986) (e.g., Fodor & Pylyshyn, 1988; Pinker & Prince, 1988) and subsequent models of IM therefore adopted alternative representational schemes for the distributed encoding of phonology. These schemes typically saw words as sequences of phonemes, with each
phoneme encoded in a distributed manner over a fixed number of phonetic and articulatory features. Distributed codes applied not only to phonemes within the stems but also to inflectional suffixes and their allomorphs (past tense allomorphs: /t/, /d/, /^d/) (e.g., Thomas & Karmiloff-Smith, 2003, though see Plunkett & Marchman 1991, 1993, 1996 for a 2-bit code for past tense suffixes).

The transition of models to the use of more plausible phonological representations was accompanied by the use layers of hidden units, either in multi-layered feed-forward architectures (MacWhinney & Leinbach, 1991; Plunkett & Marchman, 1991, 1993, 1996; Thomas, 2005; Thomas & Karmiloff-Smith, 2003) or in networks with recurrent connections and attractor structures (Hoeffner, 1992; Hoeffner & McClelland, 1993; Joanisse, 2004; Joanisse & Seidenberg, 1999). Layers of hidden units enhanced the representational capacity of neural networks, allowing them to learn mappings defined upon the new distributed codes (cf. Plunkett & Marchman, 1991). For reasons of simplicity, layers of hidden units are not included in the definition of model architectures in the last column of Table 1.

Representing lexical-semantics

Rumelhart and McClelland (1986), as well as Plunkett and Marchman (1991, 1993, 1996) trained architectures to associate mappings between phonological representations of stems and past-tense forms. Most models, however, have employed architectures that included lexical-semantic information. The inclusion of lexical-semantics information was necessary to allow models to solve the homophone problem (e.g., wring/wringed vs. ring/rang, see MacWhinney & Leinbach, 1991). Studies have also shown that lexical-semantics are particularly important for the learning and the processing of irregular mappings such that the loss of semantic
information has differential impact on irregular verb performance (e.g., Joanisse & Seidenberg, 1999).

Lexical-semantic information has been encoded either at a lemma level, i.e., through localist (orthogonal) vectors allowing the identification of individual lexical items; or at a feature level, i.e., through sparsely distributed representations of the meanings of individual lexical items, sometimes incorporating prototype structures. There appear, however, no empirical data to constrain how lexical-semantics information interacts with inflection. That is, although recent past-tense models postulate that information on individual words enters into the production of inflected words, it is not clear what the nature of lexical information must be. Thomas and Karmiloff-Smith (2003) compared a localist representational scheme for lexical-semantics to two possible distributed schemes, one based on random assignment of features to base forms, and one based on representations that were alterations of a prototype distributed structure. The localist scheme provided a better fit to behavioural data, whilst still able to capture semantic dimensions observable in the empirical data, such as, concrete vs. abstract. By contrast, rich semantic representations reduced the regularity effect in acquisition and impacted on generalization of the past-tense rule, as meaning-based rather than phonological information became increasingly important in driving performance (Thomas & Karmiloff-Smith, 2003).

Representing the linguistic environment

In the literature of connectionist models of IM, there are two main approaches to creating a training set of inflectional mappings. The first is to develop an artificial language that incorporates the main characteristics of the target domain (e.g., the
THE MULTIPLE INFLECTION GENERATOR

English past tense). The second is to derive realistic inflectional paradigms from language corpora.

The use of artificial languages to model linguistic environments was initiated in MacWhinney & Leinbach (1991) and Plunkett and Marchman (1991, 1993, 1996). Plunkett and Marchman (1991, 1993, 1996), in particular, developed an artificial vocabulary of random monosyllabic strings, generated via three templates (CVC, VVC, and CVC; C: Consonant, V: Vowel) and conforming to the phonotactics of the English language (well-formedness; see Plunkett & Marchman, 1993). These triphonemic strings were assumed to be English-like verb stems. They were inflected in the past tense taking into account constraints on the type frequencies of regular and irregular mappings or quasi-regular clusters of irregular past tenses (e.g., arbitrary: be/was, vowel change: know/knew, identity: hit/hit). Plunkett and Marchman (1991, 1993, 1996) also considered an arbitrary and highly simplified scheme of token frequency for stem/past-tense mappings distinguishing two levels, high and low.

The use of artificial linguistic environments offers the advantage of experimental control, which is particularly important for isolating the effects of psycholinguistic variables on language development. Plunkett and Marchman (1991, 1993, 1996) performed a systematic investigation of the effects of type and token frequency in the acquisition of the English past tense, as well as of the effects of the gradual exposure of the network to the full training set (incremental training). They suggested that the past-tense network underwent a transition from a stage in which morphological production behaved as if it was based on rote-learning to a stage in which it behaved as if it was rule-based, and that this transition depended on the frequencies of regular and irregular categories in the training set (Plunkett & Marchman, 1993). Further, a refined implementation of incremental training
demonstrated that overgeneralization errors were a microphenomenon characterising the acquisition of individual verb forms (Plunkett & Marchman, 1991).

An alternative approach to developing a training set is to consider natural language corpora. A number of connectionist models (e.g., Hoeffner & McClelland, 1993; Joanisse, 2004; Joanisse & Seidenberg, 1999; Woollams et al., 2009) have used existing English verbs extracted from child-directed speech or adult-language corpora. Verb stems in these models were nevertheless usually monosyllabic, i.e., similar to the artificial language approach. Variable word lengths in terms of phonemes were however possible, achieved by accommodating monosyllabic verb stems and past-tense forms in a generic slot-based template with right alignment (CCCCVCCC-VC; C: Consonant; V: Vowel; last two slots: inflectional suffixes).

Models trained on word forms derived from natural language corpora represent progression of connectionist modeling in incorporating a plausible linguistic environment. As shown in Table 1, many computational models of IM extended earlier ones by developing training sets based on natural language (e.g., Hoeffner & McClelland, 1993 vs. Hoeffner, 1992) or with larger vocabulary sizes (e.g., Woollams et al., 1999 vs. Joanisse & Seidenberg, 1999). MacWhinney and Leinbach (1991) included multisyllabic verb forms, while a recent model by Mirković et al. (2011) considered multisyllabic forms corresponding to inflections of Serbian nouns. In the latter, multisyllabic forms were produced syllable by syllable in the output layer of a four-layered network with recurrent connections within its second hidden layer (see also Bullinaria, 1995 on the alignment problem).

Despite the theoretical merit of approaches based on realistic inflectional mappings, the advantages of a highly simplified, but fully controllable, rendition of the linguistic environment are still appealing. This is especially true for exploratory
studies of novel phenomena in the acquisition of IM. In that vein, Thomas and Karmiloff-Smith (2003) and Thomas (2005) used the artificial language training set of Plunkett & Marchman (1991, 1993, 1996) to contrast theoretical positions on the acquisition of the English past tense in atypical development. These studies also relied upon a non-incremental training regime, to avoid the need of additional simulations to control for the effects of the initial composition of the training set. In Thomas, Forrester, and Ronald (2013), the same training set was used in a model investigating interactions between genetic and environmental factors in past-tense acquisition, under a population-based modeling approach, while Thomas and Knowland (in press) considered theoretical mechanisms of language delay. Studies by Thomas and colleagues have often opted for artificial-language training sets under the assumption that these offer a reasonable approximation of the linguistic environment and the opportunity to address novel phenomena such as individual differences data better through a well-defined and well-understood set of modeling constraints (cf. Thomas & Karmiloff-Smith, 2003, p.654-655).

*Multiple inflections within a grammatical class and across grammatical classes*

As shown in Table 1, the majority of connectionist models have focused on individual subdomains of IM, especially the past tense. Some models, however, have targeted the acquisition of multiple inflections, in particular multiple verb inflections (Hoeffner, 1992; Hoeffner & McClelland, 1993; MacWhinney & Leinbach, 1991), multiple noun inflections (Mirković et al., 2011), and inflections of multiple grammatical classes (Plunkett & Juola, 1999). These models have commonly
employed additional\(^1\) sources of information in their architectures to allow the system to distinguish the inflected forms that were appropriate in a given context. For example, in the models of Hoeffner (1992), Hoeffner and McClelland (1993), and MacWhinney and Leinbach (1991) for the learning of multiple inflections of English verbs, the architectures comprised information on the semantics of the targeted verb inflection (which in these models was base form, progressive, third person singular of the present tense, past tense and past participle). Plunkett and Juola (1999), who implemented a model to simulate the acquisition of verb past tense and noun plural, augmented the Plunkett and Marchman (1991, 1993, 1996) architecture with two grammatical class units, indicating whether an item should be inflected as a verb or a noun. Finally, Mirković et al. (2011) proposed a model for producing inflected forms of Serbian nouns. Serbian is a highly inflecting language and the nominal form fuses word stems with morphemes expressing the grammatical features of number, case, and gender. This study considered a composite layer of syntax units in the input layer, representing each of the three features (number, case, and gender) in a localist manner.

A primary aim of models addressing the acquisition of multiple inflections was to show that connectionist architectures are indeed able to acquire training sets encoding a larger and more complex set of inflectional mappings than individual subdomains of IM. Hoeffner (1992), Hoeffner and McClelland (1993), and MacWhinney and Leinbach (1991) used a training set of 1925 mappings, which included the five main inflections of 385 English verbs. In Plunkett and Juola (1999), the training set comprised 946 past tenses and 2280 plurals of 2626 stems (including

\(^1\) additional to phonological and lexical semantics information, usually considered in this type of neural network architecture
an overlap between grammatical classes. Finally, the training set of Mirković et al. (2011) considered 3,244 inflected forms of 407 Serbian nouns.

Of course, it is one thing to establish the learnability of larger and more complex training sets; it is another to show that the models acquire their target domains in a psycholinguistically plausible manner. For the models of Hoeffner (1992), Hoeffner and McClelland (1993), and MacWhinney and Leinbach (1991) the two principal phenomena addressed were the differences in the order of emergence of different verb inflections, and the vulnerability of these inflections in atypical language development (developmental dysphasias; Hoeffner & McClelland, 1993). Such differences depended on the type frequencies and the level of complexity of different inflection types, as well as the relevance of a phonological/perceptual deficit affecting word-final stops and fricatives to different inflection types (in Hoeffner & McClelland, 1993). Plunkett and Juola (1999) examined the acquisition profiles of individual inflections in greater detail, addressing for example how the rates of overgeneralization in the past tense and the plural depended on the training regime (incremental, non-incremental), the type frequency of inflectional categories (i.e., irregular past-tense, regular past-tense, irregular plural, and regular plural) and their overall complexity (Marchman, Plunkett, & Goodman, 1997). They also considered the extent to which the model generalized inflectional rules in cross-categorial inflections, i.e, in denominal verbs and deverbal nouns (Kim, Marcus, Pinker, Hollander, & Coppola, 1994). Finally, the Mirković et al. (2011) simulated the role of a number of psycholinguistic variables such as surface and lemma frequency or inflectional neighbourhood size in the production of inflected forms of Serbian nouns.
Models of IM in non-English languages

Although the majority of connectionist models of morphology have focused on English, some models have addressed morphological paradigms in non-English languages (Forrester & Plunkett, 1994; Hare & Elman, 1995; Hare et al., 1995; Indefrey & Goebel, 1993; Mirković et al., 2011; Nakisa & Hahn, 1996; Plunkett & Nakisa, 1997; Ruh & Westermann, 2008, 2009). We have already referred to the model of Mirković et al. (2011), which addressed the acquisition of the inflections of Serbian nouns, as a paradigm from an inflectionally rich language. The majority of connectionist models of non-English IM, however, have focused on the acquisition of the so-called minority-default domains (Forrester & Plunkett, 1994; Hare & Elman, 1995; Hare et al., 1995; Indefrey & Goebel, 1993; Nakisa & Hahn, 1996; Plunkett & Nakisa, 1997; Ruh & Westermann, 2008, 2009). Minority-default systems are inflectional paradigms which distinguish multiple conjugational classes and in which forms corresponding to a particular conjugational class of low type frequency exhibit properties of regular or default inflection, i.e., they are preferred for non-word inflection (cf. Marcus et al., 1995; Prasada & Pinker, 1993). Minority-default behaviour has been suggested to contradict connectionist accounts of IM, mainly under the assumption that the principles of associative learning should encourage productive generalization in the most frequent conjugational classes (Marcus et al., 1995). Nevertheless, connectionist studies have been successful in simulating minority-default behaviour, because this additionally depends upon the degree of phonological and semantic coherence of the default conjugational class.

A crucial question concerning the cross-linguistic generality of the connectionist framework is whether models of non-English IM were based on the same architectural assumptions and simplifications as models of English IM. In most
non-English studies in Table 1, models learnt to classify phonological forms in the conjugational categories they belonged to (Forrester & Plunkett, 1994; Hare et al., 1995; Nakisa & Hahn, 1996; Plunkett & Nakisa, 1997, Ruh & Westermann 2008, 2009). This implied using architectures in which the output layer represented inflectional class rather than phonological form. These models therefore used different architectures then those of English IM. The model of Mirković et al. (2011) represents an exception to this pattern. The architecture of this model comprised an output layer producing phonological forms of inflected words, i.e., similar to English morphology models. However, here the architecture differed from models of English IM with respect to the input layer. The input layer in the Mirković et al. (2011) model included lexical semantics and syntactic context information, whilst lacking phonological information, which has been commonly included in all other models of morphological production. It is unclear whether the model’s behaviour would be the same if it included phonology in the input.

*Note on three additional features of connectionist models of IM*

We conclude this review of connectionist models of IM by referring to three additional features. These features are not directly related to the main modeling assumptions of the MIG, but they will become highly relevant when discussing the limitations and possible extensions of our model.

The first feature concerns the extent to which it is reasonable to abstract morphological development to a learning process based on a single task, in particular a task involving the production of inflected forms from phonological, semantic, and target inflection information (an analogue of elicited production). Many models of IM have adopted this convention; other models, however, such as Joanisse (2004),
Joanisse and Seidenberg (1999), and Woollams et al. (2009), have considered multiple phases of learning, for example ‘speaking’, ‘hearing’, ‘repeating’, and ‘generating’. Each of these phases involved the learning of mappings between different pairs of representations (‘speaking’: semantics to phonology; ‘hearing’: phonology to semantics, ‘repeating’: stem phonology to stem phonology, ‘generating’: phonology and semantics to semantics) thereby contributing different types of linguistic knowledge to the system. For example, ‘repeating’ consolidated the learning of stems and consequently regular inflections (cf. Joanisse & Seidenberg, 1999). Woollams et al. (2009) used these phases to address dissociations between past-tense production in two different tasks of inflection production, from form and from meaning. The generality across tasks achieved by this model is an advantage. On the other hand, if regular verb inflection in a given model depends on repetition training (a non-inflectional task), this is an important assumption of the model that requires some empirical support.

The second feature concerns atypical language processing. Some models of IM addressed developmental or acquired deficits by implementing conditions of processing constraints at the beginning or the end of training, correspondingly (Hoeffner & McClelland, 1993; Joanisse, 2004; Joanisse & Seidenberg, 1999; Thomas, 2005; Thomas & Karmiloff-Smith, 2003; Thomas & Knowland, in press). Such models have been used for assessing theories on the aetiology of language disorders or the effects of particular types of lesions on morphological production (Thomas, Baughman, Karaminis, & Addyman, 2012).

The final feature of connectionist models of IM concerns the focus of models on neurodevelopment. Most models of Table 1 have simplified this issue and considered architectures with a constant and pre-allocated amount of computational
resources. The models of Ruh and Westerman (2008, 2009, 2012) for the English past tense and the German past participle, however, were based on constructivist neural networks. These networks emphasised the role of task-dependent neurodevelopment, in the sense that their architecture was adapted by allocating additional processing resources when and where necessary (Ruh & Westerman, 2009). It remains to be seen how fixed versus adaptive architectures differ with respect to their account of IM.

Cross-linguistic differences of English and Modern Greek with respect to IM

The principal research aim of the MIG was to apply the same connectionist architecture to two languages very different in character with respect to IM. The main cross-linguistic differences between English and Modern Greek with respect to morphology that the model focused on were as follows:

1. *English employs morphological marking for fewer grammatical categories than Modern Greek.* The English system of IM is summarised in Table 2. English presents a high and typologically rare degree of morphological simplicity (Ragnarsdottir et al., 1999, p.578) and uses morphological suffixes to mark eight grammatical categories, namely the plural and the possessive of nouns, the progressive and the third person singular (henceforth: 3rd singular) of the present tense of verbs, the past tense of verbs and the past participle of verbs, and the comparative and superlative of adjectives. By contrast, Modern Greek is a highly inflecting language that inflects most grammatical classes (six out of ten), namely articles, nouns, adjectives, pronouns, verbs, and participles (Holton, Mackridge, & Philippaki-Warburton, 2003; Triandafillidis, 1941). Nouns, adjectives, articles, pronouns, and participles follow nominal inflection and present inflected forms (or types) corresponding to different cases (nominative, genitive, accusative, vocative) of the
singular and the plural number (Triandafillidis, 1941, p.210). Verbs present types corresponding to different persons of the singular and the plural number and these types also bear morphemes marking tense, aspect, mood, and voice (Stephany, 1997, p.185). A simplified version of verb morphology in is presented in Table 3 and discussed in greater detail at the end of this section.

2. English makes extensive use of unmarked (root) forms, whereas Modern Greek completely lacks them. Many grammatical categories are not marked in English. For example, nouns do not have grammatical gender, while verbs are marked for person only in the 3rd singular. Unmarked forms of nouns, verbs, and adjectives are therefore used extensively, in all cases where a morphological suffixation rule does not apply. On the other hand, there are no root forms of nouns, adjectives, and verbs in Modern Greek (Stephany, 1997; Varlokosta, Vainikka, & Rohrbacher, 1996). Word stems are bound morphemes, i.e., they cannot stand alone as individual words, and always need to be combined with suffixes to express case (for nominal inflection), person (for verbal inflection), and number (for nominal and verbal inflections).

3. English marks words for a single grammatical category (at most), whereas Modern Greek fuses multiple inflectional morphemes in the same word forms. As shown in Table 2, the English system of IM is based on morphological suffixes. On the contrary, the system of IM in Modern Greek is synthetic and fusional (Joseph, 2008, p.486). Case, person, and number are realized by fusing the stem, i.e., the part of the word that remains the same across the different types, with suffixes (Triandafillidis, 1941, p.210). Other grammatical categories may require the use of prefixes, infixes, as well as phonologically predicted modifications of the stem and stress shift (e.g., perfective past tense of verbs; Stavrakaki & Clahsen, 2009).
4. English morphology is either fully regular or based on a dichotomy between regulars and irregulars, whereas Modern Greek morphology is based on multiple conjugational categories. In English, inflections are either fully-regular or they can be described in terms of a clear-cut dichotomy between a predominant class of regulars and a minor class of irregular examples (e.g., past tense: 160 regulars vs. 10,000 regulars; Marslen-Wilson & Tyler, 1998). This is despite the fact that regular inflections may consider allomorphic subcategories, and irregular inflections may consider quasi-regular clusters (e.g., irregular past tense; identity: set/set, vowel-change: know/knew, arbitrary: be/was). By contrast, there are multiple conjugational classes for both nominal and verbal inflections in Modern Greek (Holton et al., 2003; Stephany, 1997; Triandafillidis, 1941, Varlokosta et al., 1996). An additional source of complexity is the combination of conjugational categories corresponding to individual grammatical features. For example, verb forms are realized fusing stems corresponding to the perfective or the imperfective aspect, with suffixes for person and number, and possibly an infix for marking the past tense (e.g., Holton et al., 2003, p.108-119). As alternatives exist for all these procedures, the result is an especially complex system of verb conjugation² (Stephany, 1997, p.185).

*A simplified version of verbal morphology of Modern Greek*

---

² Stephany (1997, p.195) notes: 'There is a lack of correlation among conjugation class (two chief types), perfective active stem formation (three main types), imperfective non-past passive conjugation (five chief types), perfective passive stem formation (two types), and past participle stem formation (five types) ...'.
To exemplify the richness of the Modern Greek system of IM, we focus on verb morphology. Table 3 provides a simplified version focusing on the indicative mood of the active voice and inflections realized in single-word forms, i.e., the present tense, the imperfective past tense, and the perfective past tense. The imperfective past tense is used to denote progression, habituality, or repetition in the past, as opposed to the perfective tense, which denotes completion (Stavrakaki & Clahsen, 2009, p.115). The perfective past tense is semantically similar to the English past tense and this is possibly a reason for which it has been the focus of many psycholinguistic studies of Modern Greek.

Table 3 includes forms corresponding to all persons of the singular and the plural number of the three tenses. It distinguishes four major conjugational categories, namely 1, 2a, 2b, and 3. This taxonomy aims to accommodate conjugational classes observed within each of the three tenses. For example, the present tense (first column) considers two conjugational classes, each employing its own set of suffixes for marking person and number (Holton et al., 2003, p.119-139). The first set of suffixes applies to conjugational categories 1, 2a, and 2b, while the second to conjugational class 3.

Similarly to the present tense, the imperfective past tense also considers two basic types of inflections, the former applying to conjugational classes 1, 2a, and 2b and the latter to class 3. The first inflectional type (example of conjugational class 1: E-tre-cha = to run: imperfective-past-singular-1st person) is realized by fusing four

---

3 Similarly to other studies (e.g., Smith, 2008; Stavrakaki & Clahsen, 2009), we adopt an approximate phonetic transcription of Modern Greek based on the Latin alphabet. Words are segmented in syllables and capital letters are used to indicate vowels in stressed syllables. For the details on the phonetics of Modern Greek, see Arvaniti (2007), Holton et al. (2003), Stephany (1997).
constituents: 1) the imperfective verb stem; 2) the past-tense infix E-; 3) stress shift (from penultimate to antepenultimate) in all persons, apart from the 1st and the 2nd person of the plural, and 4) the corresponding inflectional suffixes for person and number. The second inflectional type (e.g., mi-IOU-sa, to talk: imperfective-past-singular-1st person) does not bear the infix (E-). Similar to the first type, it employs the perfective stem; however, it uses a different set of suffixes.

The last column of Table 3 describes conjugational classes within the perfective past tense. In general, the perfective past is realized similarly to the imperfective, the only difference being the use of a verb stem corresponding to the perfective instead of the imperfective aspect. The perfective stem can be produced by the addition of the aspectual marker -s- (sigma in Modern Greek) to the imperfective stem. However, alternatives exist and this yields the four conjugational categories of Table 3 (based on Ralli, 1988; Stavrakaki & Clahsen, 2009): 1) the ‘sigmatic’ class, including verbs in which the perfective stem is generated by the addition of the aspectual marker; 2a) the allomorphic class, in which the perfective stem is generated via idiosyncratic phonological modifications of the imperfective stem; 2b) the arbitrary class, in which the perfective stem is arbitrary (suppletive) and, further, the perfective past tense is formed without an infix; and 3) the class of sigmatic verbs with ultimate stress, which employ the aspectual marker, but not the past-tense infix.

**Target empirical phenomena for the acquisition of English IM**

The acquisition of English IM has been studied extensively in the literature and the available empirical data are ample. Table 4 summarises the seven key phenomena that have been observed in these data and were set as the target empirical phenomena for the MIG. Asterisks in the second column mark phenomena within English that, to our
knowledge, have not been addressed previously with computational modeling. The third column includes the studies that provided data we used for comparisons with the MIG, with font style providing a preview on how successful the model was in simulating these data (quantitative fit, qualitative fit, or dissimilar to the data; see Method for details).

*Target empirical phenomenon ENG1: Order of emergence of inflections*

Target empirical phenomenon ENG1 refers to the order in which different inflections emerge in child language. The relevant data come from the longitudinal corpus-based study of Brown (1973) and the cross-sectional study of de Villiers and de Villiers (1973). Brown (1973) analysed utterances produced by three children to compare the stages at which the rates of correct usage of different grammatical morphemes in obligatory contexts exceeded 90% for the first time (Brown’s criterion for acquisition; Brown, 1973). The progressive of verbs was acquired first, followed by the plural of nouns, the irregular past tense of verbs, and the possessive of nouns. Regular past tense and 3rd person singular were acquired later in development. De Villiers and de Villiers (1973) obtained a similar order under a cross-sectional research design.

In both Brown (1973) and de Villiers and de Villiers (1973), the order of emergence of inflections was highly correlated (rank-order correlations >0.8) with the complexity of individual inflections. The level of complexity of different inflections was the number of rules required for the derivation of morphemes according to the transformational grammar of Jacobs and Rosenbaum (1968) (cumulative syntactic complexity, cf. Brown, 1973) or the number of unitary meanings that morphemes encode in child language (cumulative semantic complexity, cf. Brown, 1973). There were, however, no reliable correlations, between morpheme frequencies in parental
speech and the order of acquisition (de Villiers and de Villiers, 1973). It is hard to explain the relationship between low-level features (complexity, frequency) of inflections and their order of acquisition under symbolic accounts of language acquisition since rules are equally complex/simple and particular grammatical categories (variables) and algebraic rules operating on them emerge in child language in an all-or-nothing fashion (cf. Marcus et al., 1999). The effect of low-level features on the order of acquisition is more consistent with the connectionist framework. For example, Plunkett and Juola (1999) who examined the acquisition of the past tense of verbs and the plural number of nouns in English found that the type frequencies of different inflection types are integrated with their level of complexity, as well as with the frequencies of the individual mappings (token frequency) to determine the order of acquisition.

The aim of the MIG with regards to target empirical phenomenon ENG1 was to be the first connectionist model to generate a rank order for the range of English inflections studied in Brown (1973) and de Villiers and de Villiers (1973). The order of emergence of inflections was based on the same criterion for acquisition (90% accuracy; Brown, 1973) and was compared to the empirical data numerically, i.e., based on the calculation of correlation coefficient values between vectors of rank orders in the model and the data. We also examined how the complexity of different inflections, as well as type frequencies derived from the tagged Brown Corpus (Francis & Kučera, 1982) and embedded in the training set and the training regime of the model related to the order of acquisition.

**Target empirical phenomena ENG2, ENG3, ENG4, and ENG5: Accuracy, error patterns, frequency effects and generalization in quasi-regular domains**
Target empirical phenomena ENG2 to ENG5 refer to developmental patterns across a range of empirical findings, such as differences in the accuracy rates in regular and irregular inflection, the occurrence and rates of particular error types (omission errors: *Yesterday, I eat a candy*; overgeneralizations: *Yesterday, I eated a candy*; and blend errors: *Yesterday, I ated a candy*), the presence of increased effects of token frequency in irregular inflection compared to regular, and the high rates of rule-based inflection of novel items. The relevant data used to assess the ability of the MIG to simulate target empirical phenomena ENG2 to ENG5 come from a past tense elicitation task considered in van der Lely and Ullman (2001). We performed qualitative and quantitative comparisons between the developmental trajectories of the MIG and three groups of typically developing children of increasing mean age (three groups of 12 children, with a mean age of 5;9, 6;11, and 7;11 years) in that study. The comparison focused on periods of the training time of the model in which the performance of the MIG matched the empirical data from the three groups of children on accuracy rates in the regular past tense.

We also considered evidence from other studies on complementary qualitative characteristics of target empirical phenomena ENG2 to ENG5. This evidence referred to the observation that the rates of blend errors are lower than the rates of overgeneralization (Marcus et al., 1992), the rates of rule-based inflection of novel items increase with their phonological similarity to existing regulars (Prasada & Pinker; 1993), and the rates of overgeneralization are higher in the plural number than the past tense (e.g., English plural: Marchman et al., 1997).

The main research aim of the MIG with regards to target empirical phenomena ENG2 and ENG5 was to examine whether qualitative and quantitative characteristics of developmental patterns in accuracy rates, frequency effects, error patterns, and the
inflection of novel items can be simulated in a neural network architecture exposed to inflectional mappings corresponding to a fully-fledged English morphological system. This issue has not been addressed under the connectionist framework and is not trivial. For example, the empirical data (e.g., van der Lely & Ullman, 2001) suggest that children do not make commission errors, i.e., they do not apply suffixes corresponding to the progressive of verbs (-ing) or the 3rd singular/noun genitive/noun plural (-s) in cases where a verb stem needs to be marked for past tense. To acquire English in a psycholinguistically plausible manner, the MIG should also not produce this error type. In a system performing similarity-based processing and exposed to a training set in which words are frequently marked with an -s suffix, such responses might well occur.

Three further phenomena were beyond the scope of the current version of the MIG for reasons of simplicity and tractability. The three limitations of the model were the following: 1) it did not address data on the U-shaped learning curve for irregulars, as empirical effects characterising the very early morphological development were beyond its scope; 2) it did not study conditions under which novel stems rhyming with existing irregular stems are inflected irregularly, i.e., similarly to their rhymes; 3) it did not address data focusing on phonological consistency, e.g., semi-regular clusters within irregular inflection, e.g., vowel-change (know/knew, grow/grew).

Target empirical phenomena ENG6: Limited effect of phonotactics

Target empirical phenomenon ENG6 refers to data from Marshall and van der Lely (2006) addressing differences between regular and irregular past tense forms in English in terms of their conformity to phonotactical rules of the language. Irregular past tenses always conform to these rules by presenting clusters of phonemes that are
also present in monomorphemic words (monomorphemically legal clusters, MLC). By contrast, the addition of a past-tense suffix to stems of regular verbs may result not only in forms with MLCs (e.g., pack/packed), but also with phonemic clusters that do not conform to the English phonotactics (monomorphemically illegal clusters, MIC; e.g., judge/judged).

When comparing accuracy rates of three groups of typically developing children of increasing chronological age in the two types of regular past-tense forms, Marshall and van der Lely (2006) found that these were higher for forms employing MLCs (conforming to the phonotactical rules) than for forms with MICs (violating phonotactical rules). However, the difference was numerical and did not reach significance levels. Marshall and van der Lely (2006) argued that this finding is easier to explain within dual-route accounts of IM rather than connectionist accounts. This was since the former predicts that symbol manipulation should not be affected by phonological characteristics, while the latter is subject to frequency and phonological similarity effects. Marshall and van der Lely (2006) went on to show that an effect of phonotactics was present, however, in cases of specific language impairment (SLI; Leonard, 1998).

The MIG is the first computational model to address this phenomenon. It did so by including in the training set a particular category of mappings corresponding to morphologically neutral (base) forms and serving to impose phonotactical constraints on inflections in the artificial language. These allowed a contrast of the acquisition of two subcategories of past-tense mappings, those with MLCs and those with MICs. The comparison between the model and the empirical data was qualitative. The magnitude of the effect was also compared with the magnitude of the standard effect of token frequency in regular inflection. On separate work we show how the presence
of an effect of phonotactics could be simulated under conditions corresponding to SLI (Karaminis, 2012; Karaminis & Thomas, in preparation).

Target empirical phenomena ENG7: Interplay between phonology and lexical semantics in derivational morphology

Target empirical phenomenon ENG7 does not refer to inflectional but to derivational morphology. It also addresses data on adult English speakers, in particular data from a study by Ramscar (2002). Although the focus of the MIG is on the acquisition of inflections, Ramscar’s (2002) data were useful to delineate the impact of different types of information in inflection.

Derivational morphology refers to the inflection of derived words, such as the derived verb ‘to brake’. This is inflected regularly in the past tense, unlike its homophone ‘break’, which is inflected irregularly. These examples imply that phonology of the root form is not the sole factor that determines whether a derived word is inflected regularly or irregularly. Ramscar (2002) performed a series of experiments showing that semantic features (meaning) also play a role in inflection. Two of these experiments were targeted by the MIG. The first experiment measured the preference for regular or irregular past tense inflection for novel verbs (e.g., frink) rhyming with both existing regulars (e.g., blink/blinked) and existing irregulars (e.g., sink/sank). Whether these verbs were inflected regularly or irregularly depended on semantics, in particular on whether the context primed the regular verbs (either rhymes or non-rhymes) or the irregular rhymes. The second experiment examined whether semantics also affected the preference for the inflection of verbs derived from nouns (denominal). It considered novel rhymes of existing regular and irregular verbs presented as denominal verbs (e.g., novel noun: a sprink, derived verb: to sprinkle, to sprink,
rhyming regular: blink, rhyming irregular: drink). Again, the preference for regular or irregular inflection depended on whether the context primed the regular and irregular rhyme.

In the MIG, we implemented an analogue of the above two experiments of Ramscar (2002) on trained versions of the model. It was not possible to achieve full correspondence with the two experiments because within the training corpus it was not possible to identify novel verbs or existing nouns that rhyme with both existing regular and irregular verbs. However, certain manipulations were appropriate for tapping the role of phonology, semantics, and derivational status in the preference for rule-based inflection. In our first simulation of Ramscar’s (2002) results, we created novel verbs rhyming with existing irregular verbs and measured whether these were inflected regularly or irregularly in the past tense in a semantic context priming an existing regular or irregular verb and a grammatical context suggesting derivation from a noun (denominal verb). In the second experiment, we measured the same preference for existing noun forms that were assigned to the grammatical class of verbs. The comparison between the model and the empirical data was qualitative, i.e., based on whether regular or irregular semantics influenced the preference of the model for rule-based inflection in the same direction as in Ramscar’s (2002) data.

Although further details on derivational morphology are beyond the scope of this paper, it is perhaps worth noting that Ramscar’s (2002) data challenged evidence provided in an earlier study by Kim, Pinker, Prince, and Prasada (1991). The latter authors suggested that derivational status, e.g., whether a novel verb has been derived from a verb (deverbal) or a noun (denominal), rather than semantics determines inflection of novel items. This claim had been based on findings that adults prefer regular or irregular inflection for irregularly-sounding novel verbs
depending on their derivational status, and that such preferences were not affected by the semantic proximity of the deriving and the derived words.

**Target empirical phenomena for Modern Greek**

The target empirical phenomena for the acquisition of Modern Greek IM are listed in Table 5. Target empirical phenomena GR1 to GR4 refer to noun morphology; target empirical phenomena GR5 to GR7 refer to adjective morphology; and target empirical phenomena GR8 to GR11 refer to verb morphology. Similarly to Table 4, Table 5 includes information on the studies that provided the empirical data for comparison (third column); whether it was possible consider quantitative comparisons between the simulation output and the data (fourth column); and a preview of the model’s successes and failures in capturing the different target empirical phenomena.

Target empirical phenomena GR1 to GR9 were addressed based on qualitative descriptions of the course of acquisition of Modern Greek, mainly from corpus-based approaches (Christofidou; 2003; Katis, 1984; Stephany, 1997; Stephany & Christodou, 2009; and Varlokonta et al., 1996). Quantitative comparisons were possible for target empirical phenomena GR10 and GR11 (data from Mastropavlou, 2007; Smith, 2008; and Stavrakaki & Claehsen, 2009). The data of Stavrakaki and Clahsen (2009) on the acquisition of the perfective past tense defined developmental trajectories for accuracy rates and error patterns in different conjugational classes. These data were used to perform comparisons parallel to those between the MIG and the data of van der Lely and Ullman (2001) on the acquisition of the English past tense.

As discussed earlier, the MIG targeted important cross-linguistic differences between English and Modern Greek with respect to morphology (Modern Greek marking more categories, lacking unmarked forms, fusing the stem with multiple
morphemes, and presenting multiple conjugational classes). To help the reader who is not familiar with Modern Greek establish a certain level of correspondence between the key target empirical phenomena in the acquisition of English and Modern Greek IM, we group the latter into four main types: 1) phenomena related to an analogue of the Optional Infinitive stage (Wexler, 1994) in Modern Greek; 2) phenomena related to the order of emergence of different grammatical features; 3) phenomena related to the developmental profile of the perfective past tense based on the sigmatic/non-sigmatic distinction; and 4) phenomena related to the effects of phonological salience in the perfective past tense.

*Target empirical phenomena related to analogues of the Optional Infinitive stage: GR1, GR5, GR8*

The absence of unmarked forms in Modern Greek implies that inflection omission errors are not possible. This is problematic for accounts of morphological development such as the Optional Infinitive (Wexler, 1994; see also Rice, Wexler, & Cleave, 1995) positing that this error type is due to certain grammatical categories (e.g., Tense and Agreement; Schütze & Wexler; 1996) missing, being underspecified, or optional in early child language. Indeed, later versions of this theory (e.g., Unique Checking Constraint, UCC; Wexler, 1999) have included modifications to address phenomena in acquisition of other languages (e.g., Danish; Wexler, 2000). In a similar vein, a number of studies on the acquisition of IM in Modern Greek aimed to identify early developmental error patterns that could be an analogue of omission errors in the productions of children acquiring English as a first language.

With regards to verbal morphology, target empirical phenomenon GR1 refers to the observation that early productions of children are characterised by the overuse
of verb forms bearing the perfective or the imperfective stem and ending in -i (i-forms, Katis, 1984; Stephany, 1997; Varlokosta et al., 1996). This could correspond to a developmental stage descriptively similar to the Optional Infinitive stage with these forms serving as the default paradigm (unmarked). Katis (1984) and Stephany (1997) proposed that the overuse of i-forms denotes that 3rd person singular forms (see Table 3, present tense), which are acquired earlier than other person/numbers, are overgeneralized in inappropriate contexts. Therefore, the overuse of i-forms corresponds to Subject-Verb agreement errors, in the sense that children fail to mark verbs in the person denoted by the subject of the sentence. Under an alternative account, i-forms correspond to the active perfect participle, a verb form without person and tense marking (Varlokosta et al., 1996).

With regards to noun inflection, target empirical phenomenon GR5 describes the overuse of noun forms ending in a vowel, which according to Stephany (1997, p.213) correspond to 'adult accusative singular forms of the three genders, as well as the nominative of neuter and feminine nouns’. Such forms have been termed as base forms (Christofidou, 2003) or all-purpose unmarked forms (Stephany & Christofidou, 2009) of nouns. Adjective inflection is similar to noun inflection apart from the fact that adjectives are also inflected with respect to gender (while nouns can be one of the following: masculine, feminine or neuter). Target empirical phenomenon GR8 refers to the overuse of neuter forms of adjectives, and in particular, nominative/accusative forms of singular number, in other contexts (Stephany, 1997, p.224).

The aim of the MIG with regards to the target empirical phenomena GR1, GR5 and GR8, was to capture early developmental error patterns in the inflection of nouns, verbs, and adjectives. Unlike the English version of the model, these phenomena were addressed in the absence of a strong prototype effect of base-form-
to-base-form mappings in the training set. This was a challenge of the model as it implied different types of error patterns for nominal and verbal inflection.

*Target empirical phenomena related to the order of emergence of grammatical features: GR2, GR3, GR4, GR6, GR7, and GR9*

Target empirical phenomena GR2, GR3, GR4, GR6, GR7, and GR9 refer to the order of emergence of the different grammatical features that the language distinguishes. For example, target empirical phenomenon GR2 states that the number of nouns emerges earlier than case in child language (Stephany, 1997). Since nouns in Modern Greek bear obligatory marking of case and number (no unmarked forms), identifying the order of acquisition of different grammatical features is based on their contrastive use (Staphany, 1997). Thus, as early forms of nouns correspond to accusative singular forms, the acquisition of case is denoted by the emergence of genitive singular, while the acquisition of number is denoted by the emergence of accusative forms of the plural number.

The MIG generated data on the order of emergence of different grammatical features in Modern Greek by adopting Brown’s (1973) criterion for the acquisition of inflections, extended from the English model. Findings from the simulation were compared qualitatively with the descriptions in the empirical data.

*Target empirical phenomenon related to developmental error patterns: GR10*

Target empirical phenomenon GR10 refers to the detailed developmental profile of the acquisition of the perfective past tense in Stavrakaki and Clahsen (2009). These authors considered the fundamental distinction between a statistically dominant class of verbs that form their past tenses based on morphological modifications according
to the so-called sigmatic rule (conjugational classes 1 and 3, in Table 3) and a less frequent class of verbs having non-sigmatic past-tense forms (conjugational classes 2a and 2b, in Table 3). They found that children’s scores in an elicited production task were higher in the sigmatic (regular) than in the non-sigmatic (irregular) category and this difference was more pronounced at earlier developmental stages. Children overapplied sigmatic rule in the non-sigmatic category but not vice versa. Incorrect responses in the non-sigmatic category were imperfective past-tense forms or perfective past tense forms of another verb. Finally, sigmatic past-tense forms were preferred for novel rhymes of both existing sigmatic and existing non-sigmatic verbs.

The MIG aimed to simulate the learning profile of the perfective past tense in Modern Greek considering quantitative comparisons with the human data, i.e., on the calculation of a correlation coefficient value for the corresponding vectors. The model and the human data were matched on performance on the sigmatic (‘regular’) category, in parallel with the comparison with English past-tense data of van der Lely and Ullman (2001). Similar comparisons between the model and the human data were performed regarding the inflection of novel items.

**Target empirical phenomenon related to phonological salience: GR11**

The last target empirical phenomenon for the MIG refers to the potential role of phonological salience in the production of the perfective past tense. The relevant data came from a study by Mastropavlou (2007) contrasting accuracy rates in verbs that employ or do not employ the syllabic augment E- (see conjugational classes 1, 2a vs. class 3 in Table 3) in their perfective past tense forms. According to Mastropavlou (2007), this distinction results in past tense forms differing in terms of phonological salience (high and low, correspondingly). Accuracy rates in these two categories were
also compared with accuracy in the arbitrary class of verbs (conjugational class 2b),
to examine the relative importance of phonological salience and regularity in
perfective past tense inflection.

Using an elicited production task, Mastropavlou (2007) found that the rates of
correct responses of children were higher in verbs forming their perfective past tense
with a syllabic infix. This could suggest the involvement of phonology in the learning
of the perfective past tense, possibly by compensating for the difficulty to acquire this
morphological rule. A corresponding pattern of higher rates of perfective past tense
forms employing an infix was also found in the preferences for the inflection of novel
rhymes of verbs in the two categories. Accuracy rates in the arbitrary class were
higher that the rates in regular verbs forming their past tenses with or without an infix.
Smith (2008) found similar patterns of increased accuracy rates in verbs of increased
phonological salience, but lower rates in verbs of the arbitrary class.

We addressed the Mastropavlou (2007) data by performing quantitative
comparisons of the model’s output and the empirical data on accuracy rates in the
perfective past tense in conjugational classes 1 and 2a (infix), 3 (no infix) and 2b
(arbitrary) of Table 3 and the preference for the inflection of rhymes of verbs in
classes 1/2a and 3.

Method

General assumptions and simplifications

Our research design entailed the development of a basic neural-network architecture
and a training procedure using training sets that reflected the main properties of the
system of IM in English and Modern Greek. Only minor adjustments were considered
for the basic architecture in the two versions of the model, accommodating cross-linguistic differences with respect to phonology.

The overarching design principles of the MIG were as follows: 1) there is an inflection system that produces inflected forms of words appropriate to the grammatical sentence context; 2) this system is responsible for producing all inflection types; 3) multiple information sources are available to drive the output of the system and therefore cues to predict the form of a given output may be exploited flexibly across development depending on the demands of particular inflection paradigms; and 4) empirical patterns in the acquisition of different languages reflect properties of the linguistic environment to which the child is exposed.

We assume that the mechanism for IM is embedded in a larger set of systems, which provide the MIG with the different types of information (see Hoeffner & McClelland, 1993). A perceptual system makes phonological representations available (e.g., Plunkett & Marchman, 1991, 1993, 1996), while a lexical knowledge system provides representations of lexical semantics (e.g., Joanisse & Seidenberg, 1999). A grammatical knowledge system contributes representations of grammatical classes (e.g., Plunkett & Juola, 1999). And a syntactic processing system signals the morphological modifications required by the context of the sentence (e.g., MacWhinney & Leinbach, 1991; Mirković et al., 2011). The acquisition of IM involves learning to integrate the multiple cues, so as to produce phonological representations of the appropriate inflected words (e.g., Thomas & Karmiloff-Smith, 2003). The output of the system of IM is propagated to the articulatory system, which produces the inflected words (Hoeffner & McClelland, 1993).

It was assumed that all types of information are well developed when the acquisition of IM commences. With respect to input and output phonology, this
assumption entailed that the child has fully developed representations of the English phonemes and the phonological form of words before learning to inflect words. With respect to lexical semantics, it was assumed that the child has fully developed representations of the meaning of individual words, or at least knowledge of individual word forms. With respect to grammatical class, it was assumed that the child knows the syntactic distinctions between different word classes, such as nouns, verbs, or adjectives. Finally, with respect to target inflection it was assumed that the child has knowledge of the semantic distinctions between different grammatical features, such as the tense of verbs, the aspect of verbs, the number of nouns, the case of nouns, or the comparison of adjectives. In many cases, these represent simplifications, as some of these sources of information have more extended developmental time courses. From an explanatory point of view, each type of information needs its own developmental account. The MIG was neutral to the details of these subsidiary accounts, though we note some debates exist. For example, Pinker (1984, 1994) proposed that grammatical categories are innate, while Schlesinger (1988) argued that they emerge from semantic categories (e.g., objects vs. action).

Assumptions and simplifications relevant to the linguistic environment

The increase in complexity of languages and inflectional paradigms occurred at the expense of some simplifications to the training sets. Following other studies, such as Plunkett and Marchman (1991, 1993, 1996), Thomas (2005), and Thomas and Karmiloff-Smith (2003), the two training sets used in the model were based on artificial languages that approximated the main phonological, morphological, and statistical characteristics of English and Modern Greek IM whilst keeping the scale of the model tractable. The MIG assumed a single phase of training referring to the
production of appropriately inflected forms, unlike models considering multiple phases of learning (‘speaking’, ‘hearing’, ‘repeating’, and ‘generating’; see Joanisse, 2004; Joanisse & Seidenberg, 1999; Woollams et al., 2009). The architecture was trained with the full set of mappings from the onset of training, in a non-incremental fashion. This simplification allowed us to avoid the need for additional simulations to control for effects of the initial composition of the training set. Incremental training has its strongest effects on the very earliest phases of development, whereas our target phenomena lay beyond this phase, where there is little difference between incremental and non-incremental training regimes.

In the English version of the model, the artificial language consisted of a vocabulary of base forms belonging to three grammatical classes (nouns, verbs, and adjectives). The training set comprised mappings describing all possible inflections for all words within each grammatical class. The mappings were constructed to reflect statistical features of English IM, including the relative frequency of grammatical classes and the frequency of allomorphic categories within inflections. These statistical features were derived from measurements on the tagged Brown corpus (Francis & Kučera, 1982), under the assumption that this collection of written documents could offer a reasonable approximation of the linguistic environment of the child (for a discussion, see Plunkett & Juola, 1999, p.467-468). More detailed accounts of morphological development should, of course, include constraints derived from child-directed corpora. The tagged Brown corpus (Francis & Kučera, 1982) was also used to derive measurements for other statistical characteristics of English, such as the frequency of inflections of nouns and verbs, or the frequency of the progressive or the past tense of verbs. These constraints were incorporated in a probabilistic training regime, which modulated the extent to which the network was exposed to
inflections of different grammatical classes and inflections within a grammatical class accordingly. Similar type-frequency schemes have also been implemented in other models considering the acquisition of multiple inflections (Hoeffner & McClelland, 1993; Plunkett & Juola, 1999; but not in Mirković et al., 2011). Finally, token frequency was considered through a highly simplified two-level scheme, involving the multiplication of weight changes calculated by the learning algorithm with two factors (1 and 3 for low and high frequency regular mappings, and 6 and 9 for irregulars, correspondingly; after Thomas & Karmiloff-Smith, 2003, see also Plaut, McClelland, Seidenberg, & Patterson, 1996).

In the Modern Greek version of the MIG, the artificial language could not include base forms as such forms do not exist in the language. For this reason, it considered stems corresponding to nouns, verbs, or adjectives. The training set consisted of stem-to-inflected-form mappings describing all the possible inflections applying to each stem. Constraints on the statistical characteristics of the system of IM in Modern Greek were obtained from measurements on the Hellenic National Corpus (Hatzigeorgiu et al., 2000) and descriptions in grammars and psycholinguistic studies (e.g., Stephany, 1997); in the absence of data, certain constraints were made parallel to the English training set. A probabilistic training regime and a simplified scheme for token frequency were implemented similarly to the English version of the model.

Finally, two additional features were included in the English training set but not in the Modern Greek version of the model. The first feature allowed the MIG to address phenomena relevant to the effects of phonotactics on past-tense inflection (Marshall & van der Lely, 2006). Phonotactical constraints were built into the linguistic environment by a special category of non-morphological mappings (base
form to base form) corresponding to words with the same word length as inflected forms (e.g., in English, inflected/non-inflected: zipped/crypt, freed/need, heated/fetid). These mappings defined a finite set of phonemic combinations that were legal as word endings. However, a subset of inflectional mappings (e.g., base form to past tense form) ended in illegal phonemic combinations, thereby violating phonotactical constraints. The second characteristic that was particular to the English version the model was the overlap between grammatical classes. It was assumed that certain base forms belonged to more than one grammatical class (e.g., ‘count’ is both a noun and a verb). The degree of overlap between the three grammatical classes was constrained based on measurements on the tagged Brown corpus (Francis & Kučera, 1982). However, we did not have any specific corpus data that would constrain the overlap between grammatical classes in the Modern Greek version of the MIG.

*Architecture*

The basic architecture used in the two versions of the MIG is depicted in Figure 1 (white area in the middle). It is a three-layered feed-forward neural network (Plunkett & Marchman, 1991, 1993, 1996; Thomas & Karmiloff-Smith; 2003) in which four types of information or cues were presented in the input layer: (Input) Phonology; Lexical Semantics; Grammatical Category; and Target Inflection. The latter indicated the type of morphological modification that the network should perform on the base form (for English) or stem (for Modern Greek) presented in the input layer of the network. The network was expected to use the four input cues to produce the phonological form corresponding to the appropriate inflected form in the output layer (Output Phonology).
Figure 1 also includes examples of input-output mappings from the English (light grey frames) and the Modern Greek (dark grey frames) training sets. In the example from English, the network produces the plural ‘cats’; in the example from Modern Greek, the network produces the 2nd person singular of the perfective past tense for the verb ‘to fall’ (E-pe-ses). These examples make reference to the representational formats considered for the different types of information employed in the architecture and illustrate the key differences between the two versions of the MIG. These issues will be addressed in further detail in the following sections. At this point, we just note that the difference between the two versions of the MIG lay in the representations for Input and Output Phonology and Target Inflection (indicated by the dotted circles in Figure 1).

**Phonological representations**

The English version of the MIG employed a distributed encoding scheme for phonemes from Thomas and Karmiloff-Smith (2003). This scheme was based on Fromkin, Blair, and Collins (2002, p.242-259) and encoded 42 phonemes, 24 consonants and 18 vowels, using 19 articulatory features. These features were the following: sonorant, consonant, syllabic, continuant, voiced, labial, anterior, +coronal, back, strident, nasal, –coronal, high, central, low, rounded, tense, and diphthong. For the purposes of the Modern Greek version model, we developed a second distributed scheme based on Arvaniti (2007). This scheme distinguished 33 phonemes, 28 consonants, and 5 vowels, based on 20 articulatory features (bilabial, labiodental, dental, alveolar, palatal, velar, plosive, affricate, nasal, trill, fricative, lateral approximant, front, central, back, close, close-mid, open, rounded and coronantal).
Similarly to Thomas & Karmiloff-Smith (2003), both base and inflected forms were encoded as sequences of phonemes, with each phoneme corresponding to a particular position (slot) of a slot-based scheme. In the English version of the model, words were monosyllabic and were accommodated in a five-slot scheme which was employed in both the input and output layer of the network (5*19=95 units; see Figure 1). The first three phonemes were accommodated in the first three slots. These phonemes could correspond to triphonemic base forms (templates: CCV, VCC, and CVC; C=Consonant; V=Vowel), irregular inflected forms (same templates as for triphonemic base forms); or word endings in the special category of base forms used to impose phonotactical constraints (templates: CCVC and CVCC). The last two slots were used to accommodate, with right alignment, inflectional suffixes or word endings of base forms constraining phonotactics.

In the Modern Greek version of the model, the slot-based scheme considered 11 slots (11*20=220 units, see Figure 1) aiming to accommodate multisyllabic words ranging from 2 to 5 syllables. Nouns, verbs, and adjectives in Modern Greek are rarely monosyllabic and bear syllabic stress in one of the last three syllables (Stephany, 1997). Syllabic stress is involved in the distinction of conjugational categories, as well as the formation of certain inflected forms (Stavrakaki & Clahsen, 2009). Based on these observations, word stems in the Modern Greek version of the MIG consisted of a full syllable and one or two consonants corresponding to the onset of a second syllable. The first syllable (templates: V, CV, and CCV) was accommodated in slots 2 to 4 and the stem ending in slots 5 and 6, both with right alignment. The first position of the slot-based scheme was used to accommodate a syllabic augment E- involved in the formation of the perfective and imperfective past tense (see Table 3), while slots 7 to 12 accommodated inflectional suffixes.
corresponding to different inflections (templates: V, VC, VCV, VCVC and VCVCVC) with right alignment. Importantly, phonological representations in the Modern Greek version of the model included three additional units to represent the syllable bearing stress, with localist encoding (e.g., 001 encoded stress on the last syllable). For Input Phonology in particular, which did not include full word forms, these units described the stress pattern of the nominative singular for nouns and adjectives, and the first person of the present tense for verbs.

Lexical-semantic representations

Lexical-semantics were represented locally, following Joannise and Seidenberg (1999) and Thomas and Karmiloff-Smith (2003). The English version of the MIG was based on a vocabulary of 1600 triphonemic and 400 four-phonemic base forms, the latter being used to constrain phonotactics. Therefore, 2,000 units of the input layer were used to encode an equal number of lemmas. The Modern Greek version of the MIG used 1,600 units for the localist encoding of 1,600 nouns, verbs, and adjectives.

Grammatical Category

Grammatical category was represented uniformly in the two versions of the MIG with three units encoding locally the membership in the grammatical class of nouns, verbs, and adjectives.

Target Inflection

Target inflection representations encoded the inflections that were possible in each of the two systems of IM. In the English version of the model, 7 units were used to encode in a localist manner 7 types of inflections: the plural number of nouns, the
possessive case of nouns, the 3rd person singular of verbs, the progressive of verbs, the past tense of verbs, the comparative of adjectives, and the superlative of adjectives. Base-form-to-base-form mappings were implemented as null inflections for all grammatical classes (all target inflection units set to zero).

In the Modern Greek version, 20 units were used to encode the targeted inflection as follows: 6 units for the localist encoding of person-number combinations (for verbs); 3 units for the localist encoding of tense (for verbs, see Table 3); 6 units for the localist encoding of case-number combinations (for nouns and adjectives); 3 units for the localist encoding of gender (for nouns); and 2 units for the localist encoding of the base or the comparative (for adjectives). Target inflection representations were thus sparsely distributed, in the sense that they concatenated several localist codes (e.g., person-number and tense for verbs).

Linguistic environment

In both the English and Modern Greek version of the model, the linguistic environment to which the architecture was exposed resulted from the combination of a training set, which included mappings describing inflections in the corresponding morphological system, and a probabilistic training regime, which ensured that the network was exposed to different inflections according to their frequency in the language. Figures 2 (English) and 3 (Modern Greek) show the structure of the linguistic environment for the two versions of the model. The coupling of the training sets with a probabilistic training is illustrated using ‘wordle’ graphs, developed using an online freeware tool (WordItOut, www.worditout.com). Wordle graphs depict the variety of types of mappings in the two training sets in the number of tags they
contain. At the same time, they depict statistical properties, with font size indicating the frequency of each inflection type (tag).

An inspection of Figures 2 and 3 reveals that the English linguistic environment presented a much simpler structure than the Modern Greek linguistic environment. Base forms, especially of nouns, were statistically dominant in the English version of the MIG (top graph in Figure 2). The middle and the bottom graphs in Figure 2 depict the quasi-regular structure of the English past tense and plural, correspondingly. The relative frequency of irregular mappings was higher in the past tense than in the plural.

The complexity of the linguistic environment in the Modern Greek version of the model (Figure 3) is reflected in an increased number of tags, compared to the English version. In the absence of default forms, differences between inflection types in terms of frequency are more even. The lower graph in Figure 3 focuses on the perfective past tense. Even when a subdomain of Modern Greek is considered individually, there is still a great deal of complexity (compare with middle graph in Figure 2), arising from the combination of different conjugational classes with different persons and numbers.

**Training Sets**

The two training sets consisted of exemplars in which input phonology, lexical semantics, grammatical class and target inflection representations mapped to output phonology. Both training sets included inflections for a vocabulary of 1600 words: 800 nouns, 400 verbs, and 400 adjectives. The distribution of words in different grammatical classes was constrained by measurements of the tagged Brown Corpus (Francis & Kučera, 1982); in the absence of relevant data the same distribution was
also used in the Modern Greek training set, since the number of nouns, verbs and adjectives is broadly constrained by the topics that people talk about. The English training set contained an additional set of 400 four-phonemic forms for constraining phonotactics, which were not inflected.

The English training set included base-form-to-base-form mappings and mappings corresponding to all inflections shown in Table 3, apart from the past participle, which was not distinguished from the past tense for reasons of simplicity. We omitted the phoneme /s/ in the –est suffix of the superlative, a simplification purely for implementation, to allow the suffix to fit in two slots. The distribution of mappings was such to include constraints on the frequencies of allomorphic categories (past tense: -/t/ : /d/ : /ed/ = 65 : 180 : 85); regular and irregular categories (past tense: 330 regulars and 70 irregulars; plural: 770 regulars and 30 irregulars; comparative and superlative: 380 regulars and 20 irregulars); and clusters within irregular mappings (e.g., irregular past tense: 50 vowel change; 10 arbitrary\(^4\); 10 identity). These constraints were based on measurements of the tagged Brown corpus (with the NLTK software, Bird, Klein, & Loper; 2009). For the full vocabulary and all inflected forms, the English training set consisted of 5,600 mappings.

The Modern Greek training set included a significantly greater degree of complexity. Verbs were inflected as shown in Table 3. Verb stems were divided in conjugational classes (150 verbs in class 1; 40 in class 2a; 10 in class 2b; 200 in class 3, based on descriptions in Stavrakaki & Clahsen, 2001) and were inflected with respect to person and number in the present tense, the imperfective past tense, and the

\(^4\)There are only two verbs with arbitrary past tenses in English. We considered a larger number of this type of mappings to allow finer graduations of performance (see also, Thomas and Karmiloff-Smith (2003, p.660).
perfective past tense. Similarly, nouns were assigned grammatical gender, divided in
conjugational classes (5 classed for masculine; 4 for feminine; and 5 for neuter), and
then inflected in the nominative, genitive, and accusative case of the singular and
plural number. Adjectives (4 classes) were inflected similarly to nouns and
additionally with respect to gender in both base and comparative. The Modern Greek
training set included 26,400 mappings, i.e., around 5 times more mappings than the
English training set.

Probabilistic training regime

The probabilistic training regime modulated the extent to which the network was
exposed to different types of inflections. For example, for mappings describing noun
inflection in English, the ratio base form : plural : genitive was set to 60 : 15 : 5
(based on measurements on the tagged Brown corpus, Francis & Kučera, 1982, using
the NLTK software; Bird et al., 2009). Similarly, in the Modern Greek version, the
frequencies of different person-number combinations and the three tenses were based
on measurements of a sample of the first 30 verbs in a randomly chosen snippet of the
HNC (e.g., 1st sing : 2nd sing : 3rd sing : 1st plur : 2nd plur : 3rd plur = 1 : 1 : 4 : 1 : 1 : 2;
present : imperfective past : perfective past = 5 : 2 : 3). Sampling of the HNC was
chosen in the absence of a tagged corpus of Modern Greek.

Generalization set

Generalization sets were developed to measure the extent to which the network was
also able to apply inflectional rules on novel items. Generalization sets included
rhymes of existing verbs, which were presented to the network with the same
grammatical class and target inflection representations but with a null lexical
semantics representation (all units set to zero). The English generalization set consisted of three subsets of novel base forms of varying degree of similarity to base forms of the training set. This was to address the effects of phonological similarity on novel-item inflection (Prasada & Pinker, 1993). In the high-similarity subset rhymes shared the last two phonemes with existing base forms; in the medium-similarity subset rhymes and existing base forms were similar only in the last phoneme; in the low-similarity subset rhymes and existing base forms shared the last phoneme, while the first two phonemes of the novel items were such that they did not follow the CVC, VCC, CVV templates used in existing items and so were phonotactically illegal. In the Modern Greek version of the model, the generalization set consisted of novel items sharing stem endings with existing stems.

It should be noted that the MIG focused on a regular generalization, i.e., we examined whether novel items were inflected similarly to existing items they rhymed with. We did not consider whether irregular rhymes were inflected irregularly. This also held in the case of derivational morphology in English model (target empirical phenomenon ENG7; data from Ramscar, 2002) where we developed a set of high-similarity irregular rhymes to examine whether these were regularised when presented in semantic contexts of regular versus irregular verbs.

Simulation design and evaluation

We performed ten replications with each version of the model, training networks that employed 75 units in the hidden layer with non-incremental training. The number of hidden units was selected based on pilot simulations with the English version of the model, as it was found sufficient to allow the network to learn all the mappings of the training set, while presenting qualitative similarities to the target empirical
phenomena for English IM in TD (cf. Thomas & Karmiloff-Smith, 2003, p.659). We assumed that the same network should also be able to acquire inflections of the Modern Greek training set.

Network weights were initialized in the interval [-1, 1] using random seeds. They were trained based on the back propagation algorithm (Rumelhart, Hinton, & Williams, 1986) with the cross-entropy learning criterion (Hinton, 1989), a pattern-update schedule, and a learning rate of 0.01. Weight changes were also multiplied by factors for token frequency (Thomas & Karmiloff-Smith, 2003). Networks were trained for 400 epochs.

In each epoch, networks were presented with a number of mappings equal to the number of the vocabulary of the artificial language (2000 for English; 1600 for Modern Greek). The two versions of the model were thus aligned in terms of their exposure to the linguistic input, to correspond to the intuition that both children acquiring English and Modern Greek as a first language are exposed a similar ‘sheer volume’ of inflectional mappings. Note, however, that this challenged the acquisition of the Modern Greek training set as in each epoch, the architecture was exposed to only ~ 6 % of its mappings, compared to ~ 35 % in the English version.

Networks were tested on the training and generalization set at the end of each epoch. For each mapping of these test sets, the output of the network was evaluated by translating the activation pattern in each slot of the output layer to a phoneme using a variant of the nearest neighbour algorithm (considering a threshold for a maximum Euclidean distance from phonemic codes; that is, a pattern was translated to the nearest legal phoneme, provided it was sufficiently close). In the English version of the MIG, the strings that were obtained by this procedure were categorised to general classes of responses based on the psycholinguistic literature and
preliminary observations of the output (e.g., past tense: correct, omission error, overgeneralization, wrong stem/correct suffix). Incorrect responses that were not captured in these categories were classified as ‘other’. In the Modern Greek version, the categorisation of output strings needed to be more fine-grained, in order to deal with the complexity and the fusional nature of the language. The defined categories described combinations of alternative responses relevant to individual features combined in a single word forms. For example, for mappings falling in the perfective past tense in conjugational class 1: Multiple error types described possible problems in the application of the sigmatic rule, combined with possibilities for errors in the suffix for person and number.

The evaluation procedure produced detailed developmental trajectories for correct responses and error patterns in different inflections, which additionally took into account fine-grained distinctions of types of mappings within a given inflection type, such as tokens of high and low-frequency, allomorphic regular paradigms, tokens of different conjugational classes, mappings conforming to phonotactics or not, and combinations of these. In this way, the output of the model was comparable to developmental data.

Qualitative comparisons identified general similarities and differences based on observations of whether the model overestimated or underestimated rates of correct responses or error patterns relevant to particular target empirical phenomena. In many cases (see Tables 4 and 5), quantitative comparisons were also possible. Such comparisons were made by calculating the Pearson’s correlation coefficient value and its significance level (two-tailed) between vectors corresponding to the model’s output and empirical data, after the model and the data were matched on certain aspects of the data (e.g., 90% accuracy for acquisition; Brown, 1973; accuracy on
regular items for the data of van der Lely & Ullman, 2001). Correlations between vectors were used because evaluation involved simultaneous comparisons between multiple measures from the model and from the empirical data. We took correlation coefficients greater than 0.8 and with a significance value less than 0.05 to imply quantitative similarities; correlation coefficients greater than 0.75 to imply qualitative similarities; otherwise the model’s output was dissimilar to empirical data. These criteria provided a strict and objective method for model-data comparison.

**Results from the English version of the MIG**

*Learnability of the English training set*

Figure 4 (continuous thick line) shows the overall accuracy of the network in the mappings of the English training set during the 400 epochs training time. Thin lines around it depict variability in accuracy rates in individual simulations. The network reached ceiling performance and overall accuracy rates exceeded 99% at the end of training. Multiple inflection types of multiple grammatical classes were therefore learnable by the neural network architecture of Figure 1. The remainder of this section examines the extent to which these inflections were also acquired in a psycholinguistically plausible manner with reference to target empirical phenomena ENG1 to ENG7.

-------------------------------------------------------------------------------

Insert Figure 4 about here

-------------------------------------------------------------------------------

*Target empirical phenomenon ENG1: Order of Emergence*
Figure 5 depicts accuracy rates for different noun, verb, and adjective inflections during the first 200 epochs of training along with Brown’s (1973) criterion for acquisition (dotted horizontal line at 90%). Table 6 focuses on a subset of inflections that were included in the studies of Brown (1973) and de Villiers and de Villiers (1973) ordered by their type frequency (second column). The third column of this table includes a simplified four-level scheme characterising inflections in terms of their morphological complexity (1: fully regular, non-allomorphic; 2: fully regular, allomorphic; 3: regular part of quasi-regular domain, allomorphic; 4: irregular), while the last three columns provide the order of acquisition in the empirical data and the MIG.

Using the same criterion for acquisition, the correlation coefficient between the rank order of acquisition in the MIG (last column of Table 6) and the rank order in Brown (1973) (fourth column) was 0.76, p=0.08; the coefficient between the rank order in the MIG and the rank order in de Villiers and de Villiers (1973) (fifth column) was 0.59, p=0.22; the coefficient between the rank orders in Brown (1973) and de Villiers and de Villiers (1973) was 0.90, p=0.01. According to the criteria for the evaluation of quantitative comparisons, the MIG fitted qualitatively the pattern of Brown (1973) but was dissimilar to the pattern of de Villiers and de Villiers (1973). The two sets of human data were, however, quantitatively similar to each other.
Qualitative similarities between the model and the empirical data were more pronounced in the acquisition of regular inflections or regular subtypes within quasi-regular inflections. The main discrepancy between the model and the human data appeared in the acquisition of the irregular past tense. This was the last to present accuracy rates over 90% in the MIG, unlike the empirical data. Arguably, this discrepancy stems from its training on the full range of the rare category of irregular mappings from the onset of the training time (not considering a set of early, mainly irregular verbs; Rumelhart & McClelland, 1986). It could be addressed in future versions of the model using incremental training regimes (see Plunkett & Juola, 1999; Plunkett & Marchman, 1993, 1996). When the irregular past tense was excluded from the comparisons between the model and empirical data, the acquisition of inflections in the MIG was quantitatively similar to the human data (correlation coefficients: 0.87, p=0.03 for Brown, 1973; 0.86, p=0.01 for de Villiers & de Villiers, 1973).

Turning to the rank order of inflections in terms of type frequencies or their ranking for morphological complexity, the rank order of type frequencies was dissimilar to the order of acquisition in both the model and the empirical data (correlation coefficients: 0.3, p=0.56 for Brown, 1973; 0.41, p=0.42 for de Villiers & de Villiers, 1973; and -0.90, p=0.01 for the MIG). The same held for the rank order of morphological complexity (correlation coefficients: 0.38, p=0.46 for Brown, 1973; 0.28, p=.06 for de Villiers & de Villiers, 1973; and 0.58, p=0.03 for the MIG). This suggests that the order of acquisition in both the MIG and the human data implied the integration of multiple statistical properties.

*Target empirical phenomenon ENG2 to ENG5: The profile of the English past tense*
Figures 6 and 7 present the learning profile of regular and irregular past tense in the empirical data from van der Lely and Ullman (2001) and the MIG. The human data provide developmental trajectories for correct responses, omission errors, and irregularised forms in regular inflection (Figure 6a), and correct responses, omission errors, and overgeneralizations in irregular inflection (Figure 7a). Figures 6b and 7b depict the output of the MIG in the regular and irregular past tense (correspondingly). The model captured the main error patterns in van der Lely and Ullman (2001), namely omission and overgeneralization errors and, similarly to the data, did not produce irregularised responses in regular inflection. The output of the model in regular inflection also included responses in which root forms were suffixed with wrong past tense allomorphs, and responses where past tense suffixes were applied to stems that were reproduced inaccurately in the output layer. It is possible that such responses are treated as correct responses in experimental tasks (experimenter perceptual biases). However, here they were classified in separate categories, namely substitution errors, and wrong stem/correct suffix errors. With regards to irregular past tense, apart from omission errors and overgeneralization, the output of the model also included blend errors. These were produced in lower rates than overgeneralization errors, in line with Marcus et al. (1992). Further, overgeneralization errors were produced in lower rates in the past tense than in noun plural (Marchman et al., 1997).
The comparison of human data and simulation results in Figure 6c (regular past tense) and 7c (irregular past tense) was performed after the model was matched to the data on accuracy in regulars. Substitution errors, wrong stem/correct suffix errors, and blend errors were excluded in the absence of evidence of how such forms were treated in van der Lely and Ullman (2001). The correlation coefficient between vectors corresponding to human performance and the modeling results, plotted in Figures 6c and 7c, was 0.96, p<0.01. Therefore, the model fitted quantitatively the data of van der Lely and Ullman (2001) (with regards to target empirical phenomena ENG2 and ENG3).

Despite the quantitative match, two limitations should be noted. First, the model produced omission errors in consistently lower rates than children, especially in regular inflection. The second limitation of the model is that irregulars were inflected less accurately than the children.

The data from van der Lely and Ullman (2001) were also used to address the interaction between token frequency and regularity across development (target empirical phenomenon ENG4). The correlation coefficient between two 12-element vectors (3 stages x 2 values for token frequency x 2 values for regularity) for accuracy rates in the MIG and the empirical data was 0.91, p<0.01, suggesting a quantitative fit. More generally, frequency-by-regularity interaction in the MIG presented three main stages (see also Ellis & Schmidt, 1998). At an early stage of language acquisition (younger group in van der Lely & Ullman, 2001), frequency effects were equally large for regular and irregular mappings. At an intermediate stage (middle and older group in van der Lely & Ullman, 2001), frequency effects were more pronounced for irregulars than for regulars. Finally, at a late stage (epoch 250 and afterwards)
accuracy rates for both regular and irregular inflections were at ceiling levels (over 95%) and frequency effects for both regular and irregulars were small.

Finally, the output of the MIG was evaluated on the inflection of novel items (target empirical phenomenon ENG5). In general, the model preferred rule-based inflection of novel rhymes and this preference was contingent on phonological similarity between novel and existing items (Prasada & Pinker, 1993). In the regular past tense, the rates of rule-based inflection (e.g., wug/wagged) at the end of training were around 84% for novel items in the high-similarity generalization subset, and 82% for items in the intermediate-similarity generalization subset. These rates were not as high (around 38%, at the end of training) for items in the low-similarity generalization subset, i.e., phonotactically illegal non-words. However, inflectional suffixes were applied. A percentage of responses were wrong stem/correct suffix errors, i.e., the correct suffix was applied to a root form that was not reproduced correctly (e.g., wug/wagged, around 41% at the end of training). Such responses were taken to signify the difficulty of the network in reproducing unusual forms than applying inflectional rules. This is a difficulty that one would expect also in children and adults akin to repeating bizarre non-words (e.g., Gallon, Harris & van der Lely, 2007). Taken together, the model responded with a regular suffix to 79% of novel items that were dissimilar to those in its training set.

Comparisons of the model output and the data of van der Lely and Ullman (2001) on the inflection of novel items were performed with the model being matched to the human data based on accuracy in existing regulars and focusing on the inflection of novel rhymes. The correlation coefficient between vectors corresponding to the matched data suggested a quantitative fit (0.85, p<0.01). However, compared to the empirical data the model produced fewer omission errors than expected.
Target empirical phenomenon ENG6: Limited effect of phonotactics

We explored whether there was an advantage in the past tense inflection of verbs conforming to the phonotactics imposed by the special category of base-form-to-base-form mappings of four-phonemic words. An advantage was observed only in the class of /t/ allomorphs. Phonotactical constraints relevant to the past tense were imposed by the presentation in every epoch of 100 base-form-to-base-form mappings corresponding to 100 of the 400 four-phonemic words (50 ending in /t/ and 50 ending in /d/). Units that represented the inflections of some words also represented the final phoneme of the base forms of other words. Under the probabilistic training regime, the word-final phonemes competed with an average of 19.5 past tense mappings ending in /-t/ and an average of 96 mappings ending in /-d/ or /-ed/. Thus, an effect of phonotactics on inflection is largely contingent on the frequencies of morphological and non-morphological exemplars. Future models could address this phenomenon in greater detail constraining these frequencies by measurements from speech and child-directed corpora.

In the simulations with the MIG reported here, the magnitude of the effect of phonotactics was relatively small. For example, the effect was weaker than the effect of token frequency (1% vs. 7% in matched intervals). A weak effect of phonotactics was consistent with Marshall and van der Lelly (2006) data on typical development. Their original data, however, also suggested a pronounced (and statistically significant) effect of phonotactics in a group of children with SLI. Although simulating SLI is beyond the scope of this paper, it is worth noting that in a series of simulations with the MIG in which the model was trained under conditions of initial processing constraints, it was possible to simulate an increased phonotactical effect.
**Target empirical phenomenon ENG7: Interplay between phonology, derivation, and semantics**

Ramscar’s (2002) data on the interplay between phonology, derivation, and semantics in the inflection of novel rhymes was addressed by designing two experiments examining the preference for regular or irregular past-tense inflection for novel denominal verbs. Our experiments could not be entirely equivalent with those of Ramscar (2002) for two reasons. First, it was not possible to identify verbs rhyming with both regulars and irregular verbs. Second, it was not possible to identify existing nouns rhyming with regular and irregular verbs. Instead, the model manipulations aimed to tap the role of lexical semantics, phonology, and grammatical class in the inflection of denominal verbs in the past tense.

In our first experiment, we created novel rhymes of irregular verbs and presented them to fully-trained networks with the grammatical class units denoting a noun (to suggest denominal derivation). The lexical semantics context comprised representations either of existing regulars or irregulars and the target inflection units were set to suggest past-tense inflection. In the second experiment, we used phonological forms of nouns presented to fully-trained networks with the grammatical class code of verbs (an alternative implementation of denominal derivation). The lexical semantics and target inflection representations were identical to the first experiment.

The results from the two experiments presented qualitative similarities to patterns in the data of Ramscar (2002). The extent to which novel items were inflected regularly or irregularly was modulated by whether the lexical semantics context was that of a regular or an irregular verb (correspondingly). In the first
experiment, when the network was presented with irregularly-sounding words carrying the derivational status of denominal verbs (an analogue of frink/drink in Ramscar, 2002), the preference for regular inflection was 71% in a regular semantic context, and 0% in the context of an existing irregular verb. In the second experiment, when the noun was transferred to the verb word class, regular inflection was 99% in a regular context, and 43% in an irregular context. By contrast, Ramscar (2002) reported modulation of regular inflection in regular vs irregular semantic context of 73% vs. 22.5% in Experiment 1 and 75% vs. 27.5% in Experiment 2.

**Results from the Modern Greek version of the MIG**

*Learnability*

Figure 4 shows overall accuracy of the MIG in the mappings of the Modern training set (thick dotted line). Thinner lines surrounding this line correspond to results from the 10 replications. The model learnt the Modern Greek training set with rates of correct responses over 97% in epoch 400. Further training for an additional interval of 100 epochs was also considered, to ensure the convergence to ceiling levels. By 500 epochs, accuracy levels had reached 99%. The learnability of the Modern Greek training set by the MIG suggested the ability of the model to acquire a notably larger and more complex training set than the English version of the model (25,600 vs. 5,600 mappings), using the same computational architecture.

Accuracy rates in the Modern Greek version of the MIG were consistently lower than accuracy rates in the English version at any given point in training. Apart from the stark contrast between the two training sets with respect to size and complexity, these differences are likely due to the alignment of the two models in terms of the sheer volume of mappings to which the two architectures were exposed.
in each epoch. Further, although the phonological representations were longer in the Modern Greek version of the MIG, the model employed the same number of hidden units. In general, the pattern of lower accuracy rates in the Modern Greek training set was not consistent with evidence from the cross-linguistic morphological acquisition. Although detailed cross-linguistic comparisons of the ages at which different inflections emerge in English and Modern Greek are beyond the scope of this paper, we can illustrate the general pattern in the cross-linguistic language development using an example from Stavrakaki and Clahsen (2009) and van der Lely and Ullman (2001). In the perfective past-tense production task considered in Stavrakaki and Clahsen (2009), rates of correct responses in the sigmatic category were over 90% at 6;4; in the English past tense production task of van der Lely and Ullman (2001), accuracy rates in regular inflection were 72.4% at 6;11.

The MIG could reproduce accuracy rates in the Modern Greek training set that were equal to or higher than corresponding rates in the English training set either by increasing the number of training experiences per epoch, or increasing the computational resources (hidden units) in the system. A greater recruitment of processing resources in response to a more complex domain could be achieved within a constructivist framework (Ruh & Westermann, 2009). Here, we can simply note that the cross-linguistic pattern for overall accuracy in the MIG suggest increased processing requirements for the acquisition of IM in Modern Greek. This prediction could be investigated using neuroimaging methodologies and constructivist artificial neural networks.

Target empirical phenomena related to analogues of the Optional Infinitive stage:

GR1, GR5, GR8
When acquiring nominal and verbal inflection, the MIG generated error patterns symptomatic of responses produced by the children during early developmental stages, associated with the inability to mark contrastively various grammatical features, such as case, person, and number (Stephany, 1997). Similar to the empirical data, these responses differed across grammatical classes and corresponded to the overgeneralization of highly frequent forms within each grammatical class to examples where other forms were appropriate. As shown in Figure 8, the acquisition of the genitive singular of neuter nouns presented high rates of forms corresponding to nominative or accusative forms of the same or other conjugational classes. As shown in Figure 9, the acquisition of verbs featured high rates of i-forms. Similar to the empirical data (Smith, 2008; Stephany, 1997; Varlokosta et al., 1996), the highest percentages of i-forms occurred in the 2nd person of the singular number, demonstrating that their occurrence was conditioned by phonological overlap with the target response.

Target empirical phenomena related to the order of emergence of grammatical features: GR2, GR3, GR4, GR6, GR7, and GR9

The MIG captured the general patterns for the order of emergence of different grammatical features described in target empirical phenomena GR2, GR3, GR4, GR6,
GR7, and GR9. Figure 10 presents results on the acquisition of the three genders of nouns, the acquisition of case and number in nouns, and the acquisition of the genitive case in the different conjugational classes of nouns (correspondingly). These patterns were identical with the relevant empirical data.

We should note, however, two important limitations of the model in addressing phenomena relevant to the order of emergence of different grammatical features. The first limitation concerned the presence of crossovers in the lines corresponding to accuracy rates in different grammatical features. One such crossover is shown in Figure 10a. Accuracy rates in feminine nouns were slightly higher than accuracy rates on neuter nouns in the early epochs of training; however, this pattern was reversed after epoch 80. Similar crossover patterns were not reported in the empirical literature. Crossovers were taken to indicate an interaction between the effects of frequency and mapping complexity in driving the behaviour of the model and the gradual acquisition of more latent regularities. In this particular case, although neuter noun mappings were more frequent, their accuracy rates were lower than accuracy rates of feminine nouns in the early epochs of training because they presented a more complex structure (e.g., four sets of plural suffixes in the neuter gender, compared to three sets in the feminine gender).

The second limitation of the MIG concerned the acquisition of the imperfective past tense. This presented higher accuracy rates than the perfective past tense and this was inconsistent with data from Stephany (1997) suggesting it is acquired late in development. This discrepancy between the modeling results and the
empirical data could be attributed to the possibility that the type frequency of the imperfective past tense was exaggerated in a training set based on a corpus of texts rather than child-directed speech, but this explanation remains to be verified.

Target empirical phenomenon related to developmental error patterns: GR10

Figures 11 and 12 compare the modeling output to the behavioural data of Stavrakaki and Clahsen (2009) for the acquisition of the perfective past tense for two main classes of verbs, the sigmatic and the non-sigtic. Empirical data came from a perfective past-tense elicitation task focusing on the 3rd person singular. The modeling output was analysed focusing on the 2nd person singular. The reason why the 2nd rather than the 3rd person singular was selected for the analysis of the simulation output was that it allowed consideration of a particular error type not presented in the 3rd person singular (see below).

There were several similarities between the simulation output and the human data with regards to accuracy rates and error patterns in the two categories of verbs. Accuracy rates were higher for sigmatic than for non-sigtic verbs, and sigmatic responses were produced in the non-sigtic category in higher percentages than non-sigtic responses in the sigmatic category. A notable percentage of responses were imperfective past tense forms in both the model and the data. The fit of the model to
the data was excellent in the sigmatic category; however, within the non-sigmatic category the model underestimated sigmatic responses and produced more ‘other’ responses.

When the model was matched to the empirical data on accuracy in the sigmatic category, the correlation coefficient between the simulation results and the data of Stavrakaki and Clahsen (2009) was 0.96, p<0.001. Therefore the model quantitatively fitted the behavioural data. In addition, a quantitative fit was also possible for the data of Stavrakaki and Clahsen (2009) for the inflection of rhymes of existing sigmatic and non-sigmatic verbs (correlation coefficient 0.94, p<0.001).

Overgeneralization of 3rd singular perfective past-tense forms. An interesting difference between the simulation results and the human data concerned the incorrect production of 3rd singular perfective past-tense forms, in the first epochs of training (Figure 11b). These forms could correspond to S-V agreement errors in the perfective past tense, i.e., responses in which the perfective past tense but not the person has been marked correctly. As S-V agreement was not considered in the perfective past-tense elicitation tasks employed in Stavrakaki and Clahsen (2009) or other group studies (e.g., Mastropavlou, 2007; Smith, 2008), the targeted empirical data did not include responses of this type. The MIG, nevertheless, predicted that this type of error should be observed in studies examining perfective past-tense formation in younger children. Another prediction was that the rates of these errors would be higher in the 2nd person singular, which presented a high degree of phonological overlap with the 3rd person singular. Although these are novel predictions of the model, the latter pattern was consistent with an analysis in a case study by Clahsen and Dalalakis
(1999) for the language of a Greek child with SLI. Further empirical evidence is warranted.

Target empirical phenomenon related to phonological salience: GR11

The quantitative comparison of the model to the data of Mastropavlou (2007) on the effects of phonological salience in the perfective past suggested that the model was disimilar to the data (correlation coefficient 0.079, p=0.9). The model presented a very small advantage in the learning of past tenses employing the syllabic infix and lower rates in the arbitrary category (note that the latter is in line with Smith, 2008; Stavrakaki & Clahsen, 2009). Addressing target phenomenon GR11 with the MIG might require a more elaborate representation of the type frequencies in Table 3. Syllabic length might also need to be considered in determining the phonological salience of different forms. The syllabic infix E- applies only to verbs with monosyllabic stems, as the perfective past tense presents forms with antepenultimate stress (Stavrakaki & Clahsen, 2009). Perfective past tense forms of verbs with multisyllabic stems do not bear an infix, but they are not necessarily forms of low-phonological salience. An account that considers syllabic length as a determinant of phonological salience challenges the taxonomy of perfective past tense forms adopted in Mastropavlou (2007).

An analysis of the emergent functional architecture of the MIG

The results from the simulations suggested that the MIG learnt training sets corresponding to fully-fledged morphological systems similar to English or Modern Greek in a way similar to the acquisition of the two languages. This was achieved through the integration of different cues in a flexible manner, i.e., with different types
of information being weighted together to determine inflection, with different cues more important for the learning of particular types of inflectional paradigms. The integration of cues was also highly contingent on the statistical characteristics of the two different linguistic environments. We investigated the progression of this process within and across the two languages by observing how the mean amplitude of weights from input units to the hidden layer, related to particular cues or mappings, changed across training time. This provides an insight into the emergence of a particular structure in the network supporting the acquisition of different inflectional paradigms. It also shows the cross-linguistic generality of the model. MIG allows for the emergence of different functional architectures for different languages. By contrast, the dual-route model (e.g., Marcus et al., 1992; Pinker, 1984, 1994, 1995, 1999; Pinker & Ullman, 2002) is only appropriate to languages presenting a dichotomy between regular and irregular inflection (e.g., English).

English version of the MIG

Figure 13 shows the progression of mean weight amplitudes corresponding to the four major cues across the training time in the English version of the MIG. Weights from target inflection units to the hidden layer had consistently the largest mean amplitude value. This pattern confirmed the obvious importance of this cue in determining inflection. Weights from input units corresponding to phonology had moderate values of mean amplitude, which were larger than the mean amplitude of weights corresponding to lexical semantics; the latter remained relatively constant across training. Of the four cues presented in the input layer, weights corresponding to the three grammatical class units had the lowest mean amplitude value. In the MIG this cue was not particularly important in inflection, verified by simulations in which
grammatical class information was omitted with no effect on developmental performance. In a sense, grammatical information was redundant, as it was encapsulated in target inflection information: when the network was asked to produce the past tense, this also implied that the item to be inflected was a verb. This also explained why denominal verbs simulated with the grammatical class information denoting the class of nouns, were readily inflected in the past tense. The two cues were incongruent, given the mappings of the training set. However, the target inflection cue was stronger and overrode the grammatical class cue.

Insert Figure 13 about here

Figure 13b presents the mean weight amplitude from the seven target inflection units for the English MIG. The lowest mean amplitudes corresponded to units encoding the plural and the genitive of nouns, as well as the 3rd singular of verbs. These three inflections shared the use of the -s suffix and its allomorphs. The -s suffix was the most common of the inflectional suffixes and applied to a wide range of regular mappings of nouns and verbs. The lower values of weight amplitudes could be due to these units being less informative than other target inflection units, in the sense that they predicted the most common morphological modification. Consistent with this observation, larger mean amplitudes were exhibited in the weights from the units encoding the comparative and the superlative of adjectives, the inflections that were less frequent in the training set. Activation of these input units needed to override more common or ‘default’ behaviour.

Finally, although the average amplitude value of weights from input units encoding lexical semantics was relatively low and constant across training, the
amplitude of these weights was highly contingent on whether these corresponded to lexical items that were regular or irregular. As shown in Figure 13c, weights from units encoding irregular items were generally stronger than weights from units encoding regular items. This difference was more pronounced within the class of verbs, possibly because the irregular cluster was more frequent within this grammatical class. The difference emerged after epoch 35, i.e., it coincided with the observation of non-zero accuracy rates in irregular mappings after this epoch (see Figure 5). Consistent with Joanisse and Seidenberg (1999), the MIG exhibited an emergent involvement of lexical semantics in irregular inflection. This was confirmed in simulations where we omitted the lexical semantics cue. The absence of lexical semantics information resulted in pronounced deficits in irregular inflection, compared to the baseline model.

The finding that weights from lexical semantics to the hidden layer were modulated by regularity suggested an emergent bipartite structure with similarities to that postulated by the dual-route model (Marcus et al., 1992; Pinker, 1984, 1994, 1995, 1999; Pinker & Ullman, 2002). The trained version of the MIG encapsulated two processes for the production of inflected forms. A regular inflection process relied heavily upon information on the phonological structure of a stem to be reproduced and combined (optionally, and as indicated by target inflection) with an appropriate suffix in the output layer of the network. An irregular inflection process, on the other hand, relied upon lexical semantics information, predicting idiosyncratic inflection for particular lexical items, and serving to block the operation of the regular process predicted by the other cues.

\footnote{Note that localist encoding allowed us to distinguish between lexical semantics units corresponding to regular and irregular items.}
Modern Greek version of the MIG

Figure 14 shows the progression of the mean weight amplitudes in the Modern Greek version of the model. Comparison of plots 14a, 14b, and 14c with the corresponding plots of Figure 13 reveals how the linguistic environment of the two versions of the model altered the emergent functional architecture. As shown in plot 14a, the mean amplitude of weights to the hidden layer coming from the target inflection was higher that the mean amplitude of weights from all other cues. This was similar to the English version of the model. However, the mean amplitude of weights from target inflection input units was lower in the Modern Greek version (mean amplitude at the end of training =1, vs 1.4 in the English version). A possible reason for this difference was the prevalence of base-form-to-base-form mappings attributing greater information content to the target inflection units.

---------------------------------------------------------------------

Insert Figure 14 about here

---------------------------------------------------------------------

Figure 14a shows mean weight amplitudes for units encoding input phonology, separating phonological information per se (articulatory features) and syllabic stress. Both parts of input phonology were important in inflection (moderate values of mean amplitude; similar to the English version). The distinction between articulation and syllabic suggested that stress information was particularly important, probably because the stress pattern underlied the assignment of lexical items to conjugational classes, and therefore determined the way items were inflected. Another important difference between the English and the Modern Greek version of the model was the high weights from units encoding grammatical class. In the English version of the
model, the information provided by grammatical class was redundant, and incorporated within target inflection. In the Modern Greek version, grammatical class information was complementary to target inflection information. For example, it was important to determine whether a given pattern for case, number, and gender referred to the inflection of a noun or adjective. Weights to the hidden layer from units encoding grammatical class were therefore stronger in the Modern Greek version of the model than the English version.

The mean amplitude of weights from target inflection units corresponding to different grammatical features was modulated by the frequency of these features in the training set. Similarly to the English version, the higher the frequency of a given grammatical feature, the lower the information content of the corresponding part of the target inflection information and the amplitude of weights from the corresponding input unit to the hidden layer. Thus, the mean amplitude of weights from units corresponding to tense is higher than the mean amplitude of weights corresponding to case and number, because the former refer only to verb mappings while the latter refer to both noun and adjective mappings (Figure 14b). In a similar manner, the mean amplitude of weights corresponding to different persons and number is higher for the first and the second person of the plural, which are less frequent. The emerging pattern is one of a system that learns ‘default’ or most frequent behaviours, and that uses strong weights to allow cues marking less frequent behaviours to override the default. Once more, the ethos of the dual-route model is present here (e.g., Marcus et al., 1992; Pinker, 1984, 1994, 1995, 1999; Pinker & Ullman, 2002). Finally, Figure 14c analyses weights from input units corresponding to different conjugational classes of verbs to the hidden layer. The amplitudes of weights presented a graded pattern that reflected the type frequencies of the different conjugational classes. This pattern
was different to the English version, due to the lack of a clear-cut dichotomy between regular and irregular inflection, reflecting the fact that a strict dual-route approach is not appropriate to highly inflected languages.

Discussion

The MIG set out to capture a wide range of empirical phenomena in the cross-linguistic acquisition of IM. We implemented a basic neural network architecture for a generalized inflectional system and exposed it to linguistic environments that incorporated the main morphological characteristics of English and Modern Greek. A principal research aim was to show that connectionist models could accommodate multiple inflectional rules (and exceptions), from multiple grammatical classes, and account for empirical effects in their acquisition. Another aim was to show that connectionist models could be general across language typologies. Developing the Modern Greek version of the MIG challenged it to acquire a system of IM with important differences from English, simulating a different range of developmental effects that describe its acquisition.

The two principal research aims of the MIG have not been previously addressed. Models of IM have been primarily focused on the English past tense. A few models that included broader inflectional paradigms have been still limited, either to the study of the acquisition of a small number (e.g., Plunkett & Juola, 1999) of inflections from different grammatical classes or the acquisition of inflections within the same grammatical class (Hoeffner, 1992; Mirković et al., 2011). These models could not offer a detailed developmental account for the emergence of fully-blown morphological systems. For example, they could not address phenomena related to the order of emergence of different inflections and simulate the main error patterns and
the corresponding rates within a general system of IM (e.g., developmental patterns in
the English past tense with the architecture being exposed to a variety of inflectional
mappings). The MIG is the first computational model to address these issues. It is also
novel in addressing data on the effect of phonotactics in inflection (Marshall & van
der Lely, 2006) and the interplay between phonology and semantics in derivational
morphology (Ramscar, 2002).

With regards to the modeling of morphological development cross-
linguistically, existing models of non-English IM have mostly focused on languages
presenting multiple conjugational classes and especially the phenomenon of minority-
default inflection. These models employed architectures that were different from
those used in studies of English IM. These architectures performed categorisation to
conjugational classes rather than inflection, or lacked phonological information in the
input layer. This was not the case for the MIG. The MIG is the first computational
model with a strong commitment to a cross-linguistic and developmental perspective,
in the sense that: 1) it employed the same architecture to address the acquisition of
different language typologies; 2) the same set of modeling assumptions and
simplifications applied to the representation formats and the development of the two
training sets; 3) the two versions were aligned with respect to their exposure to
inflectional mappings in each epoch of training time; and 4) the model was compared
to corresponding developmental data from two languages based on similar constraints
(e.g., Brown's criterion for acquisition, Brown, 1973; matching on accuracy on
regular/sigmatic past tense).

It was no small challenge to establish the learnability of training sets
corresponding to fully-fledged morphological systems in a connectionist architecture.
A greater challenge still, however, was to show that the architecture could also learn
the two training sets in a psycholinguistically plausible manner. There were numerous ways in which the model could fail. It could produce behaviours that were not symptomatic of human development. This was because the two training sets, and especially in the Modern Greek version of the model, included a rich variety of inflectional mappings that might interfere with another. Nothing in our research design and the main assumptions of the model excluded the possibility that this variation would give rise to interactions resulting in responses that were psycholinguistically unrealistic, such as, commission errors (e.g., -s suffixes in the past tense). It is therefore important that the MIG simulated the target empirical phenomena in Tables 4 and 5, as well as that in many cases the model was robust to comparisons with the empirical data under a strict numerical criteria. It is also important that the model simulated the acquisition of two different language typologies based on assumptions and simplifications that were not specific to either language.

To model the acquisition of a fully-blow morphological system, the MIG synthesised previous connectionist accounts of morphological development positing the involvement of different types of information in morphological production: phonology (Rumelhart & McClelland, 1986); lexical semantics (Joanisse & Seidenberg, 1999); grammatical class (Plunkett & Juola, 1999); and target inflection (Hoeffner, 1992). The model exemplified this multiple-cue account, showing how these four cues were integrated in a flexible manner across development to accommodate mappings from different inflections, different grammatical classes, or regular and irregular categories. It also suggested a developmental trajectory for the emergence of a structure supporting a fully-fledged system for morphological production, and demonstrated differences in this structure across languages. These
were related to major typological characteristics, such as the presence of common inflectional paradigms across grammatical classes (greater importance of the grammatical class cue in the Modern Greek version); or the presence of multiple conjugational classes (dichotomous/graded pattern for the importance of lexical semantics in the English/Modern Greek version).

Another key theoretical assumption of the MIG was the importance of statistical regularities in the linguistic input in determining developmental patterns in morphological development. Both versions of the model included psycholinguistically motivated constraints for the structure of the linguistic environment. Such constraints determined the composition of the training set and training regime. They were sufficient to drive the learning of the network in ways similar to human data, despite the simplifications of the artificial language approach and non-incremental training in the MIG. These constraints were important for explaining empirical effects in morphological development captured by the model. For example, type frequency of different inflections was integrated with complexity to determine their order of acquisition. Statistical constraints for the linguistic environment also supported a unified explanation of a range of empirical phenomena in the acquisition of English and Modern Greek. For example, omission errors in the acquisition of English, and three error patterns particular to the noun, verb, and adjective grammatical classes in the acquisition of Modern Greek, were common patterns characterising the early stages of acquisition and produced as a prototype effect of exemplars of high type frequency.

The successes of the MIG in simulating empirical effects in morphological development were not without shortcomings. For example, the English version underestimated the rates of omission errors in both the inflection of existing and novel
items. The Modern Greek version overestimated accuracy rates in the imperfective past tense. Although these shortcomings challenged the robustness of the model, they were not critical for its success in simulating the cross-linguistic morphological development, in the sense that it was possible to identify their origin in the assumptions and simplifications of the model and possible to suggest minor modifications to overcome these.

More important are, perhaps, other limitations related to major simplifications inherent in the research design of the model. These limitations need to be addressed to achieve a more plausible computational model of morphological development. For example, despite the fact that the MIG implemented a remarkably broader morphological paradigm than other models of IM, future versions of should address IM in a yet broader sense. This could include the acquisition of auxiliaries and modals or the acquisition of the noun phrase (determiner-noun) in Modern Greek. More plausible models of morphological development should also abandon the monosyllabic artificial language approach of the MIG. Such models will need to show the learnability of training sets consisting of realistic multisyllabic inflectional examples, as well as the role of constraints of the early linguistic environment of the child – derived from child-directed corpora – in empirical effects of morphological development (e.g., U-shaped learning curve for the learning of irregulars). Future versions of the MIG should also consider semantic distinctions between different words, possibly incorporated in psycholinguistically plausible distributed representations of lexical semantics (though see Thomas and Karmiloff-Smith, 2003). Such models could also include different phases of learning and differences in morphological production in different ‘modes’, e.g., inflection from stem or from meaning (Woollams et al., 2009), that is, be more general across ‘task’.
More broadly, although the focus of the MIG was on development and the extent to which changes in the learning profile of the model were similar to the profiles of children acquiring English or Modern Greek as a first language, there were several ways in which the view of morphological development in the model was static and referring to adult linguistic knowledge. For example, the model assumed a static structure of the linguistic environment (non-incremental training), a fixed amount of neurocomputational resources available to the learning system (cf. Ruh & Westermann, 2009), and that the different types of linguistic knowledge (phonology, lexical semantics, grammatical class and target inflections) are fully matured at the onset of morphological development. A fuller mechanistic account of language development will need to include developmental accounts for all these features.

Finally, it is important for a computational account of morphological development to be able to simulate deficits presented in atypical language development (e.g., SLI; Leonard, 1998). Our current work involves extending the MIG, in which we use the model to evaluate the ability of different aetiological considerations of the impairment to capture the morphological profile of SLI in English and Modern Greek.
Acknowledgement

This research was developed during the doctoral studies of the first author at Birkbeck College, London, funded by the Greek State Scholarship Foundation (IKY) and the Greek Ministry of Education. The work of the second author was supported by UK MRC Grant G0300188, the European Commission Grant NEST-029088 (ANALOGY), ESRC grant RES-062-23-2721, and a Leverhume Study Abroad Fellowship held at the University of Chicago.
References


Marslen-Wilson, W. D., & Tyler, L. K. (1998). Rules, representations, and the

doi:10.1016/S1364-6613(98)01239-X


doi:10.1093/0198250746.003.0008

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. E. (1996).
doi:10.1037/0033-295X.103.1.56

doi:10.1207/s15516709cog2304_4

doi:10.1080/016909697386691


doi:10.1017/S0140525X00043405


Ramscar, M. (2002). The role of meaning in inflection: Why the past tense
does not require a rule. *Cognitive Psychology, 45*, 45-94. doi:10.1016/S0010-0285(02)00001-4

tense morphology in Icelandic and Norwegian children: An experimental
doi:10.1017/S0305000999003918

period of Extended Optional Infinitive. *Journal of Speech and Hearing Research, 38*, 850-863.

German verb inflection. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.),
*Proceedings of the 30th Annual Conference of the Cognitive Science Society*
(pp. 2209-2216). Austin, TX: Cognitive Science Society.

Ruh, N., & Westermann, G. (2009). Simulating German verb inflection with a
constructivist neural network. In J. Mayor, N. Ruh, & K. Plunkett (Eds.),
*Connectionist Models of Behavior and Cognition II* (pp.313-324). London, UK:
World Scientific.

Representations by Error Propagation. In D. E. Rumelhart, J. L. McClelland, &
the PDP Group (Eds.), *Parallel distributed processing: Explorations in the
MA: The MIT Press.


Table 1

*Representative connectionist models of IM*

<table>
<thead>
<tr>
<th>Author(s) and year</th>
<th><strong>Scope:</strong></th>
<th><strong>Training set</strong></th>
<th><strong>Architecture and Representations</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Rumelhart &amp; McClelland, 1986</td>
<td>English, Past Tense in TD</td>
<td>420 verbs, 420 mappings</td>
<td>P&lt;sup&gt;WF&lt;/sup&gt; → P&lt;sup&gt;WF&lt;/sup&gt;</td>
</tr>
<tr>
<td>MacWhinney &amp; Leinbach, 1991</td>
<td>English, Past Tense in TD</td>
<td>2062 verbs, 6090 mappings</td>
<td>P&lt;sup&gt;D&lt;/sup&gt;, GC&lt;sup&gt;L&lt;/sup&gt; → P&lt;sup&gt;D&lt;/sup&gt; (verb morphology)</td>
</tr>
<tr>
<td>Plunkett &amp; Marchman, 1991, 1993, 1996</td>
<td>English*, Past Tense in TD</td>
<td>500 verbs, 500 mappings</td>
<td>P&lt;sup&gt;D&lt;/sup&gt; → P&lt;sup&gt;D&lt;/sup&gt;</td>
</tr>
<tr>
<td>Hoeffner, 1992</td>
<td>English*, Verb Morphology in TD</td>
<td>200 verbs, 800 mappings</td>
<td>P&lt;sup&gt;D&lt;/sup&gt;, LS&lt;sup&gt;D&lt;/sup&gt;, GC&lt;sup&gt;L&lt;/sup&gt; ↔ P&lt;sup&gt;D&lt;/sup&gt;</td>
</tr>
<tr>
<td>Hoeffner &amp; McClelland, 1993</td>
<td>English, Verb Morphology in SLI</td>
<td>385 verbs, 1925 mappings</td>
<td>P&lt;sup&gt;D&lt;/sup&gt;, LS&lt;sup&gt;D&lt;/sup&gt;, GC&lt;sup&gt;L&lt;/sup&gt; ↔ P&lt;sup&gt;D&lt;/sup&gt;</td>
</tr>
<tr>
<td>Hare, Elman &amp; Daugherty, 1995</td>
<td>Old English*, Past tense in TD (adult)</td>
<td>150 verbs, 150 mappings</td>
<td>P&lt;sup&gt;D&lt;/sup&gt;, LS&lt;sup&gt;L&lt;/sup&gt; → CC&lt;sup&gt;L&lt;/sup&gt;</td>
</tr>
<tr>
<td>Forrester &amp; Plunkett, 1994; Forrester &amp; Plunkett, 1994</td>
<td>Arabic</td>
<td>323 verbs</td>
<td>P&lt;sup&gt;D&lt;/sup&gt; → CC&lt;sup&gt;L&lt;/sup&gt;</td>
</tr>
<tr>
<td>Authors, Year</td>
<td>Language(s)</td>
<td>Task(s)</td>
<td>Data Size</td>
</tr>
<tr>
<td>---------------------</td>
<td>-------------</td>
<td>--------------------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Nakisa &amp; Hahn, 1996</td>
<td>German</td>
<td>Plural in TD</td>
<td>323 mappings</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4,273 nouns</td>
</tr>
<tr>
<td>Plunkett &amp; Nakisa, 1997</td>
<td>Arabic,</td>
<td>Plural in TD</td>
<td>323 mappings</td>
</tr>
<tr>
<td>Joanisse &amp; Seidenberg, 1999</td>
<td>English,</td>
<td>Past Tense in AD</td>
<td>600 verbs, 1200 present and past tense mappings</td>
</tr>
<tr>
<td>Plunkett &amp; Juola, 1999</td>
<td>English,</td>
<td>Past tense and Plural in TD</td>
<td>2626 stems, 2280 past tense and 946 plural mappings</td>
</tr>
<tr>
<td>Joanisse, 2004</td>
<td>English,</td>
<td>Past Tense in SLI</td>
<td>600 verbs, 1200 present and past tense mappings</td>
</tr>
<tr>
<td>Thomas &amp; Karmiloff-Smith, 2003; Thomas, 2005</td>
<td>English,</td>
<td>Past Tense in WS and SLI</td>
<td>500 verbs</td>
</tr>
<tr>
<td>Ruh &amp; Westermann, 2008, 2009</td>
<td>German,</td>
<td>Past tense in TD</td>
<td>20,000 stem/participle-class pairs</td>
</tr>
<tr>
<td>Woollams, Joanisse, &amp; Patterson, 2009</td>
<td>English,</td>
<td>Past Tense in TD</td>
<td>1923 verb/past tense mappings</td>
</tr>
<tr>
<td>Westermann &amp; Ruh, 2012</td>
<td>English,</td>
<td>Past Tense in TD</td>
<td></td>
</tr>
<tr>
<td>Past Tense in TD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mirković, Seidenberg, &amp; Joanisse, 2011</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serbian, Noun inflections</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P, LS→P</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2

*The system of IM in English.*

<table>
<thead>
<tr>
<th>Word class</th>
<th>Inflection</th>
<th>Morpheme</th>
<th>Allomorphs</th>
<th>Regular example</th>
<th>Irregular example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun</td>
<td>Plural</td>
<td>-s</td>
<td>/s/, /z/, /^z/</td>
<td>cat/cats</td>
<td>ox/oxen</td>
</tr>
<tr>
<td>Noun</td>
<td>Possessive</td>
<td>-s</td>
<td>/s/, /z/, /^z/</td>
<td>cat/cat’s</td>
<td>n/a</td>
</tr>
<tr>
<td>Verb</td>
<td>3rd singular</td>
<td>-s</td>
<td>/s/, /z/, /^z/</td>
<td>eat/eats</td>
<td>n/a</td>
</tr>
<tr>
<td>Verb</td>
<td>Past tense</td>
<td>-ed</td>
<td>/t/, /d/, /^d/</td>
<td>look/looked</td>
<td>eat/ate</td>
</tr>
<tr>
<td>Verb</td>
<td>Past participle</td>
<td>-ed</td>
<td>/t/, /d/, /^d/</td>
<td>look/looked</td>
<td>eat/eaten</td>
</tr>
<tr>
<td>Verb</td>
<td>Progressive</td>
<td>-ing</td>
<td>-</td>
<td>look/looking</td>
<td>n/a</td>
</tr>
<tr>
<td>Adjective</td>
<td>Comparative</td>
<td>-er</td>
<td>-</td>
<td>smart/smarter</td>
<td>good/better</td>
</tr>
<tr>
<td>Adjective</td>
<td>Superlative</td>
<td>-est</td>
<td>-</td>
<td>smart/smarest</td>
<td>good/best</td>
</tr>
</tbody>
</table>
Table 3

A simplified version of verbal morphology in Modern Greek (bold highlights prefixes and suffixes; underline indicates perfective stems; capital letters in examples denote stressed vowels).

<table>
<thead>
<tr>
<th>Conjugational category</th>
<th>Person and number</th>
<th>Present Tense</th>
<th>Imperfective Past Tense</th>
<th>Perfective Past Tense</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1st singular</td>
<td>trE-cho</td>
<td>E-tre-cha</td>
<td>E-tre-xa</td>
</tr>
<tr>
<td></td>
<td>2nd singular</td>
<td>trE-chis</td>
<td>E-tre-ches</td>
<td>E-tre-xes</td>
</tr>
<tr>
<td></td>
<td>3rd singular</td>
<td>trE-chi</td>
<td>E-tre-che</td>
<td>E-tre-xe</td>
</tr>
<tr>
<td></td>
<td>1st plural</td>
<td>trE-chu-me</td>
<td>trE-cha-me</td>
<td>trE-xa-me</td>
</tr>
<tr>
<td></td>
<td>2nd plural</td>
<td>trE-che-te</td>
<td>trE-cha-te</td>
<td>trE-xa-te</td>
</tr>
<tr>
<td></td>
<td>3rd plural</td>
<td>trE-chun</td>
<td>E-tre-chan</td>
<td>E-tre-xan</td>
</tr>
<tr>
<td>2a</td>
<td>1st singular</td>
<td>plE-no</td>
<td>E-ple-na</td>
<td>E-pli-na</td>
</tr>
<tr>
<td></td>
<td>2nd singular</td>
<td>plE-nis</td>
<td>E-ple-nes</td>
<td>E-pli-nes</td>
</tr>
<tr>
<td></td>
<td>3rd singular</td>
<td>plE-ni</td>
<td>E-ple-ne</td>
<td>E-pli-ne</td>
</tr>
<tr>
<td></td>
<td>1st plural</td>
<td>plE-nou-me</td>
<td>plE-na-me</td>
<td>plI-na-m e</td>
</tr>
<tr>
<td></td>
<td>2nd plural</td>
<td>plE-ne-te</td>
<td>plE-na-te</td>
<td>plI-na-te</td>
</tr>
<tr>
<td></td>
<td>3rd plural</td>
<td>plE-nun</td>
<td>E-plen-an</td>
<td>E-pli-nan</td>
</tr>
<tr>
<td>2b</td>
<td>1st singular</td>
<td>viE-po</td>
<td>E-vle-pa</td>
<td>I-da</td>
</tr>
<tr>
<td></td>
<td>2nd singular</td>
<td>viE-pis</td>
<td>E-vle-pes</td>
<td>I-des</td>
</tr>
<tr>
<td></td>
<td>3rd singular</td>
<td>viE-pi</td>
<td>E-vle-pe</td>
<td>I-de</td>
</tr>
<tr>
<td></td>
<td>1st plural</td>
<td>viE-pou-me</td>
<td>viE-pa-me</td>
<td>I-da-me</td>
</tr>
<tr>
<td></td>
<td>2nd plural</td>
<td>viE-pe-te</td>
<td>viE-pa-te</td>
<td>I-da-te</td>
</tr>
<tr>
<td></td>
<td>3rd plural</td>
<td>viE-pun</td>
<td>E-vle-pan</td>
<td>I-dan</td>
</tr>
<tr>
<td>3</td>
<td>1st singular</td>
<td>mi-iO</td>
<td>mi-iOU-sa</td>
<td>mi-li-sa</td>
</tr>
<tr>
<td>Case</td>
<td>2nd singular</td>
<td>3rd singular</td>
<td>1st plural</td>
<td>2nd plural</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------</td>
<td>--------------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>2nd singular</td>
<td>mi-lAs</td>
<td>mi-lOU-ses</td>
<td>mi-li-ses</td>
<td></td>
</tr>
<tr>
<td>3rd singular</td>
<td>mi-lA</td>
<td>mi-lOU-se</td>
<td>mi-li-se</td>
<td></td>
</tr>
<tr>
<td>1st plural</td>
<td>mi-lA-me</td>
<td>mi-lOU-sa-me</td>
<td>mi-lI-sa-me</td>
<td></td>
</tr>
<tr>
<td>2nd plural</td>
<td>mi-lA-te</td>
<td>mi-lOU-sa-te</td>
<td>mi-lI-sa-te</td>
<td></td>
</tr>
<tr>
<td>3rd plural</td>
<td>mi-lAn</td>
<td>mi-lOU-san</td>
<td>mi-lI-san</td>
<td></td>
</tr>
</tbody>
</table>
Table 4

*Target empirical phenomena in the acquisition of English IM.*

<table>
<thead>
<tr>
<th>Index</th>
<th>Phenomenon</th>
<th>Study providing data for comparison</th>
<th>Quantitative comparison possible?</th>
<th>Model fits data?</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENG1</td>
<td>Order of emergence of inflections *</td>
<td>Brown (1973); de Villiers and de Villiers (1973)</td>
<td>YES</td>
<td>YES, quantitatively</td>
</tr>
<tr>
<td>ENG2</td>
<td>Error types I and II: overgeneralization and blend errors</td>
<td>van der Lely and Ullman (2001)</td>
<td>YES</td>
<td>YES, quantitatively</td>
</tr>
<tr>
<td>ENG3</td>
<td>Error type III: Omission errors</td>
<td>van der Lely and Ullman (2001)</td>
<td>YES</td>
<td>YES, quantitatively</td>
</tr>
<tr>
<td>ENG4</td>
<td>Frequency-by-regularity interaction</td>
<td>van der Lely and Ullman (2001)</td>
<td>YES</td>
<td>YES, quantitatively</td>
</tr>
<tr>
<td>ENG5</td>
<td>Generalization</td>
<td>Prasada and Pinker (1993); van der Lely and Ullman (2001)</td>
<td>YES</td>
<td>YES, quantitatively</td>
</tr>
<tr>
<td>ENG6</td>
<td>Limited effect of phonotactics *</td>
<td>Marshall and van der Lely (2006)</td>
<td>NO</td>
<td>YES, qualitatively</td>
</tr>
<tr>
<td>ENG7</td>
<td>Derivational morphology</td>
<td>Ramscar (2002)</td>
<td>NO</td>
<td>YES, qualitatively</td>
</tr>
</tbody>
</table>
Table 5

*Target empirical phenomena in the acquisition of Modern Greek IM.*

<table>
<thead>
<tr>
<th>Index</th>
<th>Phenomenon</th>
<th>Study providing data for comparison</th>
<th>Quantitative comparison possible?</th>
<th>Model fits data?</th>
</tr>
</thead>
<tbody>
<tr>
<td>GR1</td>
<td>Accusative singular forms serving as base forms of nouns</td>
<td>Christofidou (2003); Stephany (1997); Stephany and Christofidou (2009)</td>
<td>NO</td>
<td>YES, qualitatively</td>
</tr>
<tr>
<td>GR2</td>
<td>Number and gender of nouns emerge before case</td>
<td>Christofidou (2003); Stephany (1997); Stephany and Christofidou (2009)</td>
<td>NO</td>
<td>YES, qualitatively</td>
</tr>
<tr>
<td>GR3</td>
<td>Late acquisition of the genitive case</td>
<td>Christofidou (2003); Stephany (1997); Stephany and Christofidou (2009)</td>
<td>NO</td>
<td>YES, qualitatively</td>
</tr>
<tr>
<td>GR4</td>
<td>Late acquisition of rare conjugational categories</td>
<td>Stephany (1997)</td>
<td>NO</td>
<td>YES, qualitatively</td>
</tr>
<tr>
<td>GR5</td>
<td>Accusative neuter forms serving as base form of adjectives</td>
<td>Stephany (1997)</td>
<td>NO</td>
<td>YES, qualitatively</td>
</tr>
<tr>
<td>GR6</td>
<td>Number and gender and case of adjectives are acquired similarly to number, gender, and case of nouns</td>
<td>Stephany (1997)</td>
<td>NO</td>
<td>YES, qualitatively</td>
</tr>
<tr>
<td>GR7</td>
<td>Late emergence of the comparative</td>
<td>Stephany (1997)</td>
<td>NO</td>
<td>YES, qualitatively</td>
</tr>
<tr>
<td>GR8</td>
<td>i-forms serve as base form of verbs/ Subject-Verb</td>
<td>Katis (1984); Stephany (1997); Varlokosta et al. (1996)</td>
<td>NO</td>
<td>YES, qualitatively</td>
</tr>
<tr>
<td></td>
<td>Agreement</td>
<td>Katis (1984); Stephany (1997)</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>---</td>
<td>------------------------------------------------</td>
<td>-----------------------------</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>GR9</td>
<td>Emergence of aspect and tense</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GR10</td>
<td>Perfective past tense: sigmatic vs. non-sigmatic</td>
<td>Smith (2008); Stavrakaki and Clahsen (2009)</td>
<td>YES</td>
<td>YES, qualitatively</td>
</tr>
<tr>
<td>GR11</td>
<td>Perfective past tense: phonological salience and regularity</td>
<td>Mastropavlou (2007); Smith (2008)</td>
<td>YES</td>
<td>NO</td>
</tr>
</tbody>
</table>
Table 6

Comparison of the rank order of the acquisition of inflections in Brown (1973), de Villiers & de Villiers (1973), and the MIG.

<table>
<thead>
<tr>
<th>Inflection</th>
<th>Rank order of type frequencies</th>
<th>Morphological complexity</th>
<th>Rank order in Brown (1973)</th>
<th>Rank order in de Villiers and de Villiers (1973)</th>
<th>Rank order in the MIG</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOUNS: Regular Plural</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>VERBS: Regular Past Tense</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>VERBS: Progressive</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>VERBS: 3rd Singular</td>
<td>4</td>
<td>2</td>
<td>6</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>NOUNS: Genitive</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>VERBS: Irregular Past Tense</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>6</td>
</tr>
</tbody>
</table>

Key for the third column (Morphological complexity). 1: fully regular, non-allomorphic; 2: fully regular, allomorphic; 3: regular part of quasi-regular domain, allomorphic; 4: irregular regular part of quasi-regular domain.)
**Figure captions**

**Figure 1.** The architecture of the MIG. The light grey frames analyse the structure of the input and the output representations of the network in the English version of the model; in particular, when the network is asked to produce the plural of the noun ‘cat’. The dark grey frames explain the structure of the input and output representations in the Greek version; in particular, when the network is asked to produce the 2nd person singular of the perfective past tense for the stem ‘pEft-’ (corresponding to the Modern Greek verb for the meaning ‘to fall’). The two main cross-linguistic differences in the application of the architecture to the two languages concern: (1) the inclusion of an increased number of phonemes, as well as stress, in the phonological representations of the Modern Greek version; and (2) the structure of the Target Inflection representations, which include inflections/grammatical categories appropriate to the two languages.

**Figure 2.** Graphical depiction (‘wordle’ graphs) of relative frequencies of different items in the English training set. Frequencies were based on a tagged corpus and larger fonts indicate higher type frequencies. The top graph (enclosed in the rounded rectangle drawn with a solid line) depicts type frequencies in the whole training set. The middle and bottom graphs (enclosed in rounded rectangles drawn with a dashed and a dotted line) refine type frequencies in the past tense of verbs and the plural of nouns (respectively), distinguishing allomorphs and clusters of irregulars. The middle and the bottom graphs are mapped onto the corresponding elements of the top graph.

**Figure 3.** Graphical depiction (‘wordle’ graphs) of relative frequencies of different items in the Modern Greek training set. Frequencies were based on sampling of the
HNC corpus (Hatzigeorgiu et al., 2000) and descriptions in Stephany (1997) and Stavrakaki and Claessen (2009). Larger fonts are used to indicate higher type frequencies. The top graph (enclosed in the rounded rectangle drawn with a solid line) depicts type frequencies in the whole training set. The bottom graph (rounded rectangle drawn with a dashed line) focused on the perfective past tense, including conjugational classes and different person/number combinations. The bottom graph is mapped onto the corresponding elements of the top graph. There is a many-to-one correspondence (unlike Figure 1) indicative of the fusional character of Modern Greek IM.

Figure 4. Overall accuracy of the MIG in the English and the Modern Greek training set. The thick lines (continuous: English; dotted: Modern Greek) are the average performance of the model over 10 replications. The thin coloured lines depict performance in 10 replications. In each epoch of training the network was exposed to a number of mappings equal to the size of the vocabulary of the artificial language, i.e., 2000 for English, 1600 four Modern Greek.

Figure 5. Accuracy rates for different inflections in the English version of the MIG for the first 200 epochs of training. The black horizontal line at 90% corresponds to the criterion for the order of emergence of inflections considered in Brown (1973).

Figure 6. (a) Learning profile of the regular past tense in van der Lely and Ulman (2001); (b) Learning profile of the regular past tense in MIG; (c) Model output versus human data on regular past tense. The comparison of the model with the human data is based on three stages of training (first: epochs 18-32, second: epochs 43-63, and
third: epochs 63-79) in which the model and the human data were matched on correct performance on regular verbs.

**Figure 7.** (a) Learning profile of the irregular past tense in van der Lely and Ulman (2001); (b) Learning profile of the irregular past tense in MIG; (c) Model output versus human data on irregular past tense. The model and the human data are matched on correct performance on regular verbs (three stages of training: epochs 18-32, 43-63, and 63-79; see also Figure 6).

**Figure 8.** Error patterns in the genitive case of the singular number for different conjugational categories of neuter nouns. Continuous lines indicate overgeneralizations of nominative/accusative (‘default’) forms, while dashed lines indicate overgeneralizations of genitive suffixes from other conjugational classes. Brackets in the legend of the panel include more detailed descriptions of individual error patterns. For example, ‘acc-sing’ means that these forms are overgeneralizations of accusative singular forms of the same conjugational class, while ‘gen-sing like masc 2a, 2b’ suggests that these forms are overgeneralizations of the genitive singular forms of the conjugational classes 2a and 2b of masculine nouns.

**Figure 9.** First and second person singular of the present tense of verbs (conjugational class 1a): Accuracy rates (continuous lines) and error patterns (dashed and dotted lines). Error patterns suggest that the MIG captures the production of i-forms, i.e., an analogue of ‘default’ inflection in the grammatical class of verbs.
Figure 10. (a) Accuracy rates in masculine, feminine, and neuter nouns. Neuter nouns reached accuracy rates of 90% first (Brown's criterion; Brown, 1973), followed by feminine, and then neuter nouns; (b) Accuracy rates in the nominative plural and the genitive singular of nouns. The nominative plural was acquired earlier suggesting that number emerged earlier than case (nominative and accusative singular are identical for the majority of the nouns); (c) Rates of correct responses in the genitive singular case for different conjugational categories of neuter nouns.

Figure 11. The learning profile of sigmatic perfective past tense in the MIG compared with data from Stavrakaki and Clahsen (2009). (a) Data from Stavrakaki and Clahsen (2009) on sigmatic verbs; (b) The learning profile of the 2nd person singular of the conjugational class 1a; (c) Human data versus modeling results, for sigmatic verbs. Comparisons were based on matching the model and the human data (six age groups) on performance on sigmatic verbs (area after the vertical line in plot b).

Figure 12. The learning profile of the non-sigmatic perfective past tense in the MIG compared with data from Stavrakaki and Clahsen (2009). (a) Data from Stavrakaki and Clahsen (2009); (b) The learning profile of the 2nd person singular of the conjugational class 1b1; (c) Human data versus modeling results, for non-sigmatic verbs. The model and the human data were matched on performance on sigmatic verbs (area after the vertical line in plot e, cf. Figure 11).

Figure 13. Mean amplitude of weights from the input to the hidden layer across the training time for the English version of the MIG. (a) Weights corresponding to parts of the network encoding the four basic types of information presented at the input layer, i.e., phonology, lexical semantics, grammatical class, and target inflection; (b)
Weights from the units that encode different inflections; (c) Weights corresponding to the semantics of regular and irregular nouns, verbs, and adjectives, and weights from the units encoding the three grammatical classes. (Note that training time increases non-linearly in the horizontal axis).

Figure 14. Mean amplitude of weights from the input to the hidden layer across the training time for the Modern Greek version of the MIG. (a) Weights corresponding to parts of the network encoding the four basic types of information presented at the input layer, i.e., phonology, lexical semantics, grammatical class, and target inflection. Unlike, the English version, phonology in the Modern Greek version of the MIG also includes syllabic stress; the mean weight from units encoding it are depicted separately; (b) Weights from target inflection units encoding different grammatical categories (thicker lines), as well as person-number combinations; (c) Weights from units encoding the semantics of verbs of the four conjugational classes (continues lines), and weights from the units encoding the three grammatical classes. (Note that training time increases non-linearly in the horizontal axis).
Figures

Figure 1

\[
\begin{array}{|c|c|c|c|}
\hline
\text{GR} & \text{ENG} & \text{Output Phonology} & \text{Hidden Layer} \\
\hline
\text{GR} & \text{ENG} & \text{Lexical Semantics} & \text{Stem Phonology} & \text{Gram. Class} & \text{Target Inflection} \\
\hline
\end{array}
\]

\[
\begin{array}{|c|c|c|c|}
\hline
00..010...00 & /cl/ /a/ /d/ _ _ & 010 & 0100000000 \\
\text{ENG (cat - semantics)} & 5*19=75 \text{ bits} & \text{(noun)} & \text{plural} \\
\hline
00..010...00 & _ /p/ /e/ /l/ /d/ _ _ _ _ _ + 010 & 010 & 0000100000000000 \\
\text{(pEft: to fall ([stem]) pEft: to fall [stem], penultimate stress} & \text{perfecive past tense, 2nd person singular} & \text{(verb)} & \text{2nd person singular} \\
\hline
00..010...00 & 12*18=123 \text{ bits} + 3 \text{ bits stress} & & \\
\text{GR} & & & \\
\hline
\end{array}
\]
Figure 4
Figure 5

Accuracy in different inflections

- VERBS: Progressive
- NOUNS: Regular Plural
- NOUNS: Genitive
- VERBS: 3rd Singular
- ADJECTIVES: Regular Comparative
- VERBS: Regular Past Tense
- VERBS: Irregular Past Tense
- NOUNS: Irregular Plural

% correct

epochs
Figure 6
THE MULTIPLE INFLECTION GENERATOR

Figure 7

(b) HUMAN DATA: Irregular past tense

Correct
Omission errors
Overgeneralisation errors

5:9 6:11 7:11 chronological age

(a) MODEL: Irregular past tense

Correct
Omission error
Overgeneralisation error
Bash error

epochs

(c) HUMAN DATA vs MODEL: Irregular past tense

5:9 yrs
6:11 yrs
7:11 yrs

Mean of Human Data
Mean of Model

response type

corr. omiss. overg. corr. overg. irreg. corr. omiss. overg.
Figure 8

Error patterns in the genitive singular (neuter nouns)

- neut 1a: mI-lo (acc-sing)
- neut 1b: ne-r0 (acc-sing)
- neut 2: pe-di (nom-sing)
- neut 2: pe-dO (gen-sing like neut-1a,1b)
- neut 3: pA-cho (acc-sing like neut-1a,b)
- neut 3: pA-chou (gen-sing like neut-1a,1b or masc-2a,2b)
- neut 4: kI-ma (acc-sing)
Figure 9

Accuracy rates and error patterns in first and second person singular of present tense (class 1e)
Figure 10

(a) neuter > feminine > masculine

(b) number > case

(c) accuracy in the genitive singular (neuter nouns)

Graphs showing the percentage of correct predictions for different categories over epochs.
Figure 11

(a) HUMAN DATA: Non-Sigmatic past tense

(b) MODEL: Non-sigmatic past tense

(c) HUMAN DATA vs MODEL: Non-sigmatic past tense

response type
Figure 12
Figure 13
Figure 14