Lexical self-organization in the acquisition and processing of Modern Greek conjugation: an artificial neural network approach

by Stavros A. Bompolas
Postgraduate Student Number: 160

Patras
June, 2017
Lexical self-organization in the acquisition and processing of Modern Greek conjugation: an artificial neural network approach

by Stavros A. Bompolas

Supervisory Committee
Supervisor: Prof. Angela Ralli (University of Patras)

Examining Committee
Dr. Vito Pirrelli (Institute for Computational Linguistics «A. Zampolli» - Cnr)
Ass. Prof. Kyriakos Sgarbas (University of Patras)
Στον πατέρα μου
Summary

The traditional dichotomy between lexicon and grammar has been reviewed (Aitchison 1994), since human lexical knowledge does not seem to be organized in such a way to minimize the storage, but rather to maximize the processing efficiency (Post et al. 2008). These claims have led to the perception of the morphological lexicon as a dynamic memory system (Elman 1995), a parallel network of fully memorized word forms (Bybee 1995), whose boundaries between processing and storage are blurred and overlapped, being actually two sides of the same coin. Hence, the emergence of the morphological structure in the mental lexicon could be perceived through the analysis of the dynamic interaction between lexical representations and distribution capacity, perceptual discrimination and degrees of regularity in input data.

Artificial neural networks and self-organising maps provide experimental ways of testing this interaction under different input conditions and parameters (Pirrelli 2007). In this line, Temporal Self-Organising Maps (TSOMs, Ferro et al. 2011; Marzi et al. 2014, 2015, 2016) can perceive possible surface relations between word forms and store them by partially overlapping activation patterns, reflecting gradient levels of lexical specificity (holistic vs. decompositional representations).

For this dissertation, we use the above assumptions in order to observe the dynamic acquisition and processing of words grouped into families (e.g. a paradigm, or a set of forms filling the same paradigm cell) based on simulations with artificial neural networks, tested on samples of realistically distributed training data of the Greek language. To address the issues of developmental acquisition, serial processing and lexical access, as well as parallel activation and word recall, we had to run a series of simulations, designed to investigate the interconnection between time of acquisition, frequency distribution, word length and regularity of inflectional paradigms (Pirrelli et al. 2011; Marzi et al. 2014, 2015, 2016).

Simulations show that different Greek verb classes are processed and acquired differentially, as a function of their degrees of formal transparency and predictability. We relate these results to psycholinguistic evidence of Modern Greek word processing, and interpret our findings as supporting a view of the mental lexicon as an emergent integrative system.

Keywords: Standard Modern Greek, inflectional paradigm, inflectional regularity, temporal self-organising maps, word processing, morphological structure
Περίληψη

Η παραδοσιακή διάκριση μεταξύ λεξικού και γραμματικής έχει αναθεωρηθεί (Aitchison 1994), καθώς η ανθρώπινη λεξική γνώση δεν φαίνεται να οργανώνεται με τέτοιο τρόπο ώστε να ελαχιστοποιείται η αποθήκευση, αλλά να μεγιστοποιείται η αποτελεσματικότητα της επεξεργασίας (Post et al. 2008). Οι ισχυρισμοί αυτού έχουν οδηγήσει στην πρόσληψη του μορφολογικού λεξικού ως ενός συστήματος δυναμικής μνήμης (Elman 1995), ενός παράλληλου δικτύου με πλήρες αποθηκευμένες λέξεις (Bybee 1995), στο οποίο τα όρια μεταξύ επεξεργασίας και αποθήκευσης είναι θολά και αλληλεπικαλύπτονται, αποτελώντας ουσιαστικά δύο όψεις του ίδιου νομίσματος. Έτσι, η ανάδυση της μορφολογικής δομής στο Νοητικό Λεξικό θα μπορούσε να προσληφθεί μέσω της ανάλυσης της δυναμικής αλληλεπίδρασης μεταξύ των λεξικών αναπαραστάσεων και της δυναμικής κατανομής τους, των αντιληπτικών διακρίσεων και των διαφορετικών βαθμών μορφολογικής ομαλότητας στα δεδομένα εισόδου.

Τα τεχνητά νευρονικά δίκτυα και οι χάρτες αυτο-οργάνωσης παρέχουν πειραματικούς τρόπους για να εξεταστεί αυτή η αλληλεπίδραση υπό διαφορετικές συνθήκες και παραμέτρους (Pirrelli 2007). Σε αυτή τη γραμμή, το νευροϋπολογιστικό μοντέλο Temporal Self-Organising Maps (TSOMs, Ferro et al. 2011, Marzi et al. 2014, 2015, 2016) μπορεί να προσλάβει πιθανές επιφανειακές σχέσεις μεταξύ λεξικών μορφών και να τις αποθηκεύσει με μερικώς αλληλεπικαλύπτομενα μοτίβα ενεργοποίησης, αντικατοπτρίζοντας διαφορετικούς βαθμούς λεξικής εξειδίκευσης και πρόσβασης (ωλιστικές vs. αποσυνθέτικες λεξικές αναπαραστάσεις).

Με βάση τις παραπάνω θεωρητικές παραδοχές, μελετάμε τη δυναμική κατάκτηση και επεξεργασία λέξεων που ομαδοποιούνται σε οικογένειες λέξεων (δηλ. ένα κλιτικό παράδειγμα) με βάση προσομοιώσεις που γίνονται με τη χρήση τεχνητών νευρονικών δικτύων που έχουν εκπαιδευθεί σε δείγμα εισόδων από τη Νοητική Νεοελληνική γλώσσα. Προσεγγίζουμε τη σχέση μεταξύ χρόνου κατάκτησης, κατανομής συχνότητας και ομαλότητας των κλιτικών παραδειγμάτων (Pirrelli et al. 2011, Marzi et al. 2014, 2015, 2016).

Οι προσομοιώσεις δείχνουν ότι οι διαφορετικές τάξεις ρημάτων της Κοινής Νεοελληνικής κατακτώνται και υφίστανται επεξεργασία με διαφορετικό τρόπο, ανάλογα με τους βαθμούς διαφάνειας και προβλεψιμότητας της επιφανειακής δομής. Επιχειρούμε να συνεχίσουμε αυτά τα αποτελέσματα με ψυχογλωσσολογικές αποδείξεις και να ερμηνεύουμε τα συμπεράσματά μας, υποστηρίζοντας την άποψη ότι το Νοητικό Λεξικό αποτελεί ένα αναδυόμενο ενοποιημένο σύστημα.

Λέξεις-κλειδιά: Κοινή Νεοελληνική, κλιτικό παράδειγμα, τεχνητά νευρονικά δίκτυα, μορφολογική ομαλότητα, λεξική επεξεργασία, μορφολογική δομή, Νοητικό Λεξικό
Acknowledgements

In this part of my dissertation, I want to express my appreciation to several people who have been very supportive during all these years of my studies.

First of all, I owe a great debt to my supervisor, Prof. Angela Ralli. She has been very supportive since my first steps in Linguistics and, over all these years, she has trusted me and given me many opportunities to discover and test my emergent academic and research skills. For me, she is and she will be a point of reference to my life course.

Moreover, I would like to thank the second (unofficial) supervisor of my dissertation, Dr. Vito Pirrelli, who is also one of the members of the examining committee. I had the opportunity to work with him during Erasmus+ Program in 2016 and, since then, we have enjoyed an ongoing collaboration. I am very grateful as he always makes very insightful comments on my drafts and he offers me moral support as well. Furthermore, he and his research team (Claudia Marzi, Marcello Ferro, and Franco Alberto Cardillo) offered me great hospitality during my extended visit to Pisa, at the Institute for Computational Linguistics «A. Zampolli» - CNR. The present dissertation is the result of my productive collaboration with them. I am grateful to them for their invaluable contribution to the design of the experiments and the analysis of all data presented here.

I am also grateful to all the academic staff of the specialization of Linguistics at the University of Patras who have been very supportive during the two years of the postgraduate program: “Modern Approaches to Language and Texts”. I would like to thank Prof. Argyris Archakis, Prof. Anna Rousou, Assoc. Prof. George Xydopoulos, Assoc. Prof. Dimitris Papazachariou and Assist. Prof. Theodoros Markopoulos for the excellent collaboration.

I have the opportunity to work at the Laboratory of Modern Greek Dialects (University of Patras) and I am happy that I made some new friends. First of all, I have to thank one of the “older” members, Dr. Nikos Koutsoukos, for teaching me so many things during all these years. Of course, I would also like to thank the other members of the Laboratory, Dr. Dimitra Melissaropoulou, Dr. Christos Papanagiotou, Vaso Alexelli, George Chairetakis, and Michalis Marinis. I am indebted to Vasiliki Makri and Vasiliki Mouchtouri for bearing with me and for the good times we spend in the lab and out of the lab.

Last, but not least, I am very grateful to my family and my friends. I owe them a few words in Greek:

Η ύπαρξη και η ζωή μου δεν θα ορίζοταν με τον ίδιο τρόπο χωρίς την οικογένειά μου και τους φίλους μου, τους οποίους θεωρώ, επίσης, οικογένειά μου. Θέλω να ευχαριστήσω θερμά τους σωπηρούς φίλους της ζωής μου που έχουν πάντα εκεί, ακόμα και όταν δεν τους βλέπω, για να με διευκολύνουν να προσεγγίζω του στόχου μου.
# Table of contents

Summary ........................................................................................................................................... iv

Περίληψη ........................................................................................................................................... v

Acknowledgements .......................................................................................................................... vi

Table of contents ............................................................................................................................. vii

List of Abbreviations ....................................................................................................................... viii

List of Tables and Figures ................................................................................................................ ix

1. Introduction .................................................................................................................................. 1
  1.1 Aims and scope of the thesis ....................................................................................................... 2
  1.2 Organization of the thesis ........................................................................................................... 4

2. Theoretical background ............................................................................................................... 5
  2.1 Linguistic paradigms .................................................................................................................. 5
  2.2 The status of paradigms in inflectional morphology .................................................................... 5
  2.3 The status of paradigms in paradigm-based approaches to inflectional morphology ............... 8

3. Paradigm structure of MG conjugation ....................................................................................... 10
  3.1 Stem alternation related to inflection ....................................................................................... 11
  3.2 Inflectional affixes .................................................................................................................... 14
  3.3 The morphological organization of MG conjugation ................................................................. 15
  3.4 Morphological regularity in MG conjugation ........................................................................... 18

4. Word learning and processing: psycho-linguistic approaches ..................................................... 20
  4.1 Single- vs. dual-route models ..................................................................................................... 20
  4.2 Morphological (ir)regularity: experimental evidence from MG conjugation ......................... 22
  4.3 Storage vs. processing .............................................................................................................. 25

5. A neuro-computational approach to MG conjugation ............................................................... 26
  5.1 Temporal Self-Organising Maps ............................................................................................... 28
  5.2 Experimental evidence ............................................................................................................. 29
     5.2.1 Dataset and training regimes ............................................................................................ 29
     5.2.2 Word and paradigm acquisition ....................................................................................... 31
     5.2.3 Lexical access and word recall ....................................................................................... 37

6. General discussion and conclusions ........................................................................................... 40

References ........................................................................................................................................ 43
List of Abbreviations

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMU(s)</td>
<td>Best Matching Unit(s)</td>
</tr>
<tr>
<td>DoP</td>
<td>Difficulty of Prediction</td>
</tr>
<tr>
<td>DoR</td>
<td>Difficulty of Recall</td>
</tr>
<tr>
<td>I</td>
<td>Irregular(s)</td>
</tr>
<tr>
<td>IAP(s)</td>
<td>Integrated Activation Pattern(s)</td>
</tr>
<tr>
<td>IC(s)</td>
<td>Inflectional Class(es)</td>
</tr>
<tr>
<td>IR</td>
<td>Inflectional Regularity</td>
</tr>
<tr>
<td>IRType</td>
<td>Type of Inflectional Regularity</td>
</tr>
<tr>
<td>learnEp</td>
<td>learning Epoch(s)</td>
</tr>
<tr>
<td>len</td>
<td>word length</td>
</tr>
<tr>
<td>lenSt</td>
<td>length of Stem</td>
</tr>
<tr>
<td>lenSx</td>
<td>length of Suffix</td>
</tr>
<tr>
<td>LME</td>
<td>Linear Mixed Effects</td>
</tr>
<tr>
<td>logFreq</td>
<td>log Frequency</td>
</tr>
<tr>
<td>MG</td>
<td>Modern Greek</td>
</tr>
<tr>
<td>R</td>
<td>Regular(s)</td>
</tr>
<tr>
<td>RMA</td>
<td>Regular(s) with Morphologically-conditioned Allomorphy</td>
</tr>
<tr>
<td>RPA</td>
<td>Regular(s) with Phonologically-conditioned Allomorphy</td>
</tr>
<tr>
<td>SD</td>
<td>Skewed Distribution(s)</td>
</tr>
<tr>
<td>SOA</td>
<td>Stimulus-Onset Asynchrony</td>
</tr>
<tr>
<td>TSOM(s)</td>
<td>Temporal Self-Organising Map(s)</td>
</tr>
<tr>
<td>UD</td>
<td>Uniform Distribution(s)</td>
</tr>
<tr>
<td>WP</td>
<td>Word-and-Paradigm</td>
</tr>
</tbody>
</table>
List of Tables and Figures

**Table 1:** Cases of stem alternation in MG ................................................................. 12
**Table 2:** Distribution of morphologically-conditioned stem change in active voice, indicative .... 17
**Table 3:** Inflectional Class 1 (stem paradigm: ðen ~ ðe ‘tie’) ........................................ 17
**Table 4:** Inflectional Class 2a (stem paradigm ayap(a) ~ ayapi ‘love’) ................................. 17
**Table 5:** Inflectional Class 2b (stem paradigm ðier ~ ðiere ‘divide’) ................................... 18
**Table 6:** Stem alternation and (ir)regularity ..................................................................... 19

**Figure 1:** Outline architecture of a TSOM ..................................................................... 28
**Figure 2:** Average frequency of correctly recalled words by learning epochs for the SD regime.. 31
**Figure 3:** The time course of lexical acquisition (recall accuracy) on uniform maps vs. skewed maps for both type counts and token counts ...................................................... 32
**Figure 4:** Acquisition time of paradigms ranked by increasing learning epoch for the SD and UD regimes ........................................................................................................ 33
**Figure 5:** Acquisition time of regular and irregular paradigms ranked by increasing learning epoch for the SD and UD regimes ............................................................... 34
**Figure 6:** Acquisition time of regulars with morphologically-conditioned stem alternants and regulars with phonologically-conditioned stem alternants ............................................... 35
**Figure 7:** Marginal plots of interaction effects between word length, log frequency and inflectional regularity in a LME model fitting word learning epochs in MG ........................................ 36
**Figure 8:** Marginal plot of interaction effects between log frequency, length and inflectional regularity in a LME model fitting word learning epochs by TSOM instances trained on Greek verb forms in the SD regime ............................................................. 37
**Figure 9:** Marginal plot of interaction effects between log frequency, stem length, suffix length and inflectional regularity in a LME model fitting DoP by TSOM instances trained on Greek verb forms in the SD regime ................................................................. 38
**Figure 10:** Marginal plot of interaction effects between word frequency, word length and inflectional regularity in a LME model fitting DoR by TSOM instances trained on Greek verb forms in the SD regime .................................................. 39
**Figure 11:** Marginal plot of interaction effects between length and degrees of stem regularity in a LME model fitting DoR by TSOM instances trained on Greek verb forms ....................... 40
1. Introduction

How do children acquire language? In language acquisition, the most prominent datum is the human ability to deduce grammatical relations which are not learned or taught explicitly. Furthermore, it is a given that children acquire language despite the so-called poverty of stimulus. In the literature, there are several accounts for this ability which can be grouped into the traditional behavioral and linguistic approaches as well as the interactionist position of language acquisition. The contradistinction of the above positions could be perceived on the basis of their approaches to the divisions of structuralism vs. functionalism, performance vs. competence, and nativism vs. empiricism in language acquisition (Gleason & Ratner 2016: 158-194).

Behavioral and linguistic approaches to language acquisition and processing opposed since very early (Chomsky vs. Skinner debate) and their principles constitute the bases of modern theories and models. Linguistic approaches typically assume that language has a structure or a grammar which is independent of language use to some extent, determining what is grammatical (or permissible) in any given language. According to Chomsky (1957), a grammar of this type should define the speakers’ knowledge of all grammatically acceptable utterances (linguistic competence) rather than what a speaker actually utters (linguistic performance). Grammars are generative in the sense that they are comprised of a finite set of rules, common to all speakers of a given language, thus allowing the production of an infinite set of mutually understandable sentences. As a result, speakers of any language can produce and comprehend sentences they have never said or heard before, merely by making use of grammatical rules and insertion of various lexical items. Consequently, in a linguistic account, the human ability to acquire and process a language is based on an innate grammar which is realized with specific parameters of a given language (Chomsky 1965).

On the other hand, behavioral theories are characterized by an emphasis on the evident and measurable features of linguistic performance. Behaviorists reject the existence of internal structures (or processes) with no exact physical association or observable behavior, such as grammars (Skinner 1957; Zimmerman & Whitehurst 1979). Considering that mental processes are not easily observed and defined, behaviorists claim that we can obtain only implicit knowledge of the grammatical rules and structures. In contrast, observable context conditions (stimuli) that co-occur and expect precise behaviors (responses) should be looked for. Moreover, although behaviorists acknowledge that humans are physiologically developed to speak, they favor language as a skill which is not principally different from any other behavior (Watson 1924; Skinner 1957). Hence, the development of language does not lie in the knowledge of language-specific rules, but in relating various stimuli to internal responses, and these internal responses to be expressed as verbal behaviors guided by imitation, reinforcement, and successive approximations to adult performance; that is, in empiricism. In that context, the child is perceived as a passive recipient of
environmental influences in addition to speakers being considered as having no active role in the process of language development or behavior (Skinner 1957).

Turning to the interactionist approaches, they might be considered as a conciliation of the classic behavioral and linguistic approaches. These approaches recognize the strong arguments of the earlier theoretical camps, adopting assumptions from each one without prerequisites and restrictions of simple associations or strong innate devices. This can be easily deduced when the various interactive approaches are taken into consideration; the social approach, the gestural and usage-based approach, and the cognitive approach. The several interactionist approaches share the common idea that several factors (e.g. linguistic, social, cognitive, biological) are interdependent, interact with, and modify each other, affecting the progress of language development. More specifically, the social approach (e.g. Bates & MacWhinney 1982; Ninio & Snow 1999) points out the social environment in which a language is acquired and the critical aspects of social interaction, without which language acquisition appears to be impossible and maybe needless. The usage and gestural approach (e.g. Rizzolatti & Arbib 1998; Stakoe 2001; Tomasello 2003, 2008; Corballis 2010) gives prominence to both developing cognition and social interaction within which children, like early humans, begin interacting by using simple gestural communication followed by more complex linguistic forms, which arise only later after the automatization of the simpler ones. Finally, according to the cognitive interactionist approach (e.g. Piaget 1954), language is perceived as only one of the several complex cognitive skills that children acquire. In this framework, the linguistic structure and the processes implicated in its acquisition are constrained by the nature of the child’s mental status during the course of acquisition. Furthermore, in the same framework, the information-processing paradigm (Rumelhart & McClelland 1987; Bates & MacWhinney 1989; MacWhinney 1989, 1999) highlights the processing demands of language acquisition, investigating the availability and reliability of linguistic cues that indicate significant functions. In this line, the course of language development is determined by the nature of the information to be processed.

The heterogeneity of these three interactional approaches renders them the most promising for the future. The fact that they recognize the strong points of the earlier theoretical frameworks and, at the same time, they avoid a compulsive preoccupation on associative or innate mechanisms, puts an advantage on interactionists who might surpass the more obvious drawbacks resulting from solid linguistic and behavioral theories. However, until a general learning and processing model of language, that precisely specifies both the required psychological and environmental variables, is incorporated, there is a crucial need for reviewing and testing the given theoretical approaches by using new methodological tools and methods.

1.1 Aims and scope of the thesis

The present thesis does not aim to develop a new theory about morphological acquisition and processing, but to put under investigation a more recent cognitive
approach, assuming that language comprehension as well as production involve a network of parallel processing units—functionally equivalent to neuron clusters—which are connected in such a way that multiple operations or decisions proceed concurrently, selectively firing in response to sensory stimuli (e.g. McClelland & Elman 1986; Norris 1994; Luce & Pisoni 1998). Hence, this thesis has two major aims which could be generally formulated as such: First, it tries to examine the adequacy of such a neuro-computational model when it has to deal with the peculiar characteristics of Modern Greek (MG) conjugation. Secondly, a neuro-computational model of this kind provides the opportunity to investigate the descriptive and explanatory adequacy of Ralli’s (1988, 2005, 2007) seminal work on the morphological system of MG conjugation and to test traditional hypotheses of grammar and lexicon interaction in word processing and learning (one-route vs. dual-route mechanisms) along with exploring more assiduously the potential of single, distributed mechanisms in addressing word processing challenges (Alegre & Gordon 1999; Baayen 2007).

As a result, the above investigation is built up by developing both a bottom-up and a top-down approach. In the first case, the input consists of realistically distributed training data exported by a huge corpus and, at the same time, the neuro-computational approach used here allows us to build up a bottom-up investigation with a view to understanding the cognitive mechanisms that govern word (and paradigm) acquisition and processing. In the latter case, a number of theoretical questions related to the topic of MG conjugation are the building blocks for the data analysis, presetting the training regimes defined for the acquisition of the data by our model and the specific factors on the basis of which we generated the marginal plots.

This dissertation draws to a great extent on crucial insights enclosed in inspiring contributions in the literature. The present thesis largely uses as foundation Ralli’s (1988, 2005, 2007) influential work on the morphological system of MG conjugation. Ralli’s analysis is of great value because it both offers substantial support to the position that inflection should not be accounted for in a post-syntactic component, drawing data from MG, and it also presents the structural characteristics of the overall MG morphological system. In this regard, it forms the foundations of a great part of the analysis provided here. Moreover, the computational investigation of the MG verb system was carried out here with a particular family of artificial neural networks, named Temporal Self-Organising Maps (TSOMs) (Ferro et al. 2011; Marzi et al. 2014, 2015, 2016; Pirrelli et al. 2015). Unlike traditional multi-layered perceptrons, TSOMs simulate the dynamic spatial and temporal organization of memory nodes supporting the processing of time-series of symbols, to allow monitoring the short-term and long-term processing behaviour of a serial memory exposed to an increasingly larger set of word forms. TSOMs are ideal tools for assessing the differential impact of several aspects of morphological properties on the behaviour of a connectionist framework.

In conclusion, the present thesis bases its claims on a paradigm-based approach (Matthews 1972; Zwicky 1985; Anderson 1992; Stump 2001; among others) to word processing/learning, assuming that word forms are not acquired in isolation, but through associative relations, linking members of the same word family (e.g. a paradigm, or a set of forms filling the same paradigm cell). Principles of correlative
learning offer a set of dynamic equations that is the key to modelling this complex dynamic at a considerable level of detail. We use these dynamic equations to simulate acquisition of MG conjugation and we try to relate these results to psycholinguistic evidence of MG word processing, and interpret our findings as supporting a view of the mental lexicon as an emergent integrative system, as the interactionists would claim (Corbett & Fraser 1993; Bybee 1995; Wunderlich 1996; Alegre & Gordon 1999; Dressler et al. 2006; Baayen 2007).

1.2 Organization of the thesis

This thesis is structured as follows: I start with the presentation of the main theoretical models of language acquisition (and processing) (Chapter 1), as well as the main theoretical approaches of the organization of morphological components (Chapter 2), focusing on an interactionist model and a paradigm-based approach, respectively. Afterwards (Chapter 3), the paradigm structure of Modern Greek conjugation is presented on the basis of Ralli’s analysis followed by experimental data supporting her view of paradigm classes in Modern Greek (Chapter 4). At the rest of the dissertation, we account for the same evidence running a series of simulations with artificial neural networks in order to observe the dynamic acquisition and processing of words grouped into families (Chapter 5), and, finally, some conclusions are drawn (Chapter 6).

More particularly, among the different approaches to language acquisition and processing, the interactional ones are regarded as the most promising for the future paving the way for the investigation of this dissertation (Chapter 1). Therefore, after the definition of the concept of paradigms (Chapter 2.1) and their significance among the different theoretical approaches to inflectional morphology (Chapter 2.2), a paradigm-based approach is adopted, according to which the lexicon of a given language is a form-based (self-)organized network in which lexical structures are generated, as a consequence of the paradigmatic relations among fully inflected word forms (Chapter 2.2). Next, the descriptive analysis of MG conjugation is provided in detail, focusing on stem allomorphy (Chapter 3.1), affixation (Chapter 3.2), and inflectional classes (ICs) (Chapter 3.3). The seminal work of Ralli offers evidence which appears to question a dichotomous view of regularity vs. irregularity as a function of storage vs. rule-based processing mechanisms, because of the presence of a category of verb class that combines both of these mechanisms (stored-allomorphy + rule-based suffixation) (Chapter 3.4). Then, after an introduction to the two main theoretical camps of word processing, represented by single- and dual- route models of word learning and processing (Chapter 4.1), we give some experimental evidence from MG conjugation. Indeed, the experimental analysis resulting from Greek data bears confirmation to the idea of degrees or gradations of morphological regularity, as a connectionist single-mechanism would predict, rather than qualitatively different operations based on the distinction between regulars and irregulars, as the dual-mechanism approach would predict (Chapter 4.3). The hypothesis that this evidence is compatible with a parallel processing architecture (a Temporal Self-organising Map), where processing and storage are in fact mutually implied, is tested (Chapter 5.1) by running simulations on samples of realistically distributed data of the Greek language (Chapter 5.2.1). To address the issues of developmental acquisition (Chapter 5.2.2), serial processing and lexical access as well as parallel activation and word recall (Chapter 5.2.3), we had to
run a series of simulations, designed to investigate the interconnection between time of acquisition, frequency distribution, word length and regularity of inflectional paradigms. Finally, we discuss the given results and draw some conclusions (Chapter 6).

2. Theoretical background

2.1 Linguistic paradigms

Linguistic units are inherently multi-dimensional. They exhibit a variety of morphological, orthographical, phonological, pragmatic, semantic and syntactic properties. Accordingly, paradigmatic (substitutional and associative in the Saussurian sense) relations among the dimensions of a unit with other units can be established at all levels of language, e.g. the selection of /r-/ as opposed to /v-/ or /p-/ etc., in the context /-olos/ (e.g. ‘rolos ‘role’), or of to ‘the’ as opposed to ‘ena ‘a’, ‘liyo ‘a little’, ‘poli ‘much’, etc., in the context __ne ‘ro ‘water’ (Crystal 2008: 348-349). Paradigmatic relations together with syntagmatic relations constitute the statement of a linguistic unit within the language system. In that general sense, the term paradigm is used to indicate a set of linguistic units with a common property (e.g. to ‘the’ and ‘ena ‘a’ belong to the paradigm of determiners of MG) (Booij 2007: 8).

Even if paradigmatic relations cross-cut all levels of linguistic analysis, the concept of paradigm is generally restricted to the domain of morphology and, especially, to the field of inflectional morphology. The (ir)regularities associated with words grouped into paradigms has a long and notable tradition starting from classical Greek, Latin and Sanskrit grammarians. Today, a central topic in morphological theory is whether inflectional paradigms have theoretical significance (paradigm-based theories of inflection: e.g. Stump 2001, 2016; Blevins 2003, 2006; Ackerman et al. 2009; Brown & Hippisley 2012) or not (morpheme-based theories of inflection: e.g. Halle & Marantz 1993; Müller 2002; Bobaljik 2002).

2.2 The status of paradigms in inflectional morphology

The morphological process of word formation that does not change grammatical category and does not create new lexemes but rather changes the grammatical forms of a lexeme so that they fit into different grammatical contexts is called inflection. Inflectional morphology is the domain of morphology that deals with sets of inflected word forms, by investigating the relations holding within and across word paradigms (Matthews 1991: 38, 42). According to a typical analysis of Matthews (1991: 28-40, 42-54), an inflectional paradigm is the set of all grammatical words of a given lexeme. In any language exhibiting inflection, each grammatical word in a sentence carries a set of morphosyntactic properties, which are usually associated with specific aspects of its morphology. The grammatical words of a lexeme are said to be members of its paradigm. A paradigm is structured in such a way that, for each form of a given lexeme,
there exists at least one combination of morphological features that the form realizes. Each such a well-formed array of features is referred to as the cell of a paradigm.

The concept of paradigms is central in most recent theories of morphology and has traditionally played a prominent role in the study of language as one of the fundamental organizational units of the lexicon (e.g. De Saussure 1916; Zwicky 1985; Anderson 1992; Aronoff 1994; Stump 2001; Blevins 2003; among others). Nonetheless, there are some accounts of word-structure in which the paradigm, as a theoretical construct, is an epiphenomenon rather than a basic unit of morphological organization (Williams 1981; Lieber 1992; Halle & Marantz 1993; among others). Stump (2001: 1-3) makes a fourfold distinction among theories of inflectional morphology to point out that the status of paradigms in word structure and morphological organization depends on the way a theory analyses inflection. Following the literature, his divisional approach focuses on:

(i) lexical-incremental theories (e.g. Lieber 1992);
(ii) lexical-realizational theories (e.g. Halle and Marantz 1993: Distributed Morphology);
(iii) inferential-incremental theories (e.g. Steele 1995: Articulated Morphology);

In a lexical theory, the correlations between sets of morphosyntactic properties of an inflected word and its morphology are listed in the lexicon. Thus, the associations between inflectional markings and the sets of morphosyntactic properties are depicted as the association between the base form(s) of a lexeme and its grammatical and semantic properties. In a theory of this sort, since the inflectional morphemes are entries in the lexicon, they underlie the same lexical insertion principles as typical lexical morphemes. For instance, the Greek verb form ‘peksame ‘we played’ arises through the insertion of the lexically listed morphemes pek-, -s and -ame into a particular constituent structure.

On the contrary, in inferential theories, inflectional markings are not listed in the lexicon, but the systematic formal correlations between the base form(s) of a lexeme and the inflectionally complex word forms constituting its paradigm rely on rules or formulas. Hence, morphological rules infer fully inflected word forms from more basic stems (stem-based inferences) or from other word forms (word-based inferences). Each rule marks a specific step in the inference of a complex word form. For instance, ‘peksame ‘we played’ may be inferred from more basic stems through a chain of inferences: pek- → peks- → peksame or, alternatively, from the contrasting word form ‘epeksa ‘I played’, which implies (and is implied by) ‘peksame. Here, the inflectional suffixes -s and -ame are not inserted by the lexicon, but morphological rules associate their appearance with the set of the properties they present.

Stump, by intersecting this distinction between lexical and inferential theories, assumes a second distinction between incremental and realizational theories. According to incremental theories, inflectional morphology is information-increasing.
Specifically, inflectionally complex words acquire their morphosyntactic properties in steps as a consequence of acquiring their exponents. Each inflectional morpheme is associated with a particular morphosyntactic content (in the lexicon or in a rule of inference) and the acquisition occurs through the complex combination of individual inflectional morphemes which express particular morphosyntactic properties. For instance, in a theory of this kind, \( \text{peksame} \) acquires the properties preterite, first person and plural through the lexical insertion of the suffix -ame or by means of a rule inferring \( \text{peksame} \), either from the perfective stem peks- or from a related word form, e.g. \( \text{epeksa} \).

On the other hand, according to realizational theories, the association of a word with a particular set of morphosyntactic properties logically precedes and at the same time licences the realization of those properties by specific inflectional exponents. That is, this association defines the lexical insertion of its affixes (lexical theories) or the rules by which it is inferred from a stem or related word form (inferential theories). As an example, the association \( \langle \text{PEZO}, \{1 \text{ pl perfective past indicative active}\} \rangle \) licenses either the lexical insertion of the morphemes pek-, -s and -ame, or the stem-based chain of inferences pek- \( \rightarrow \) peks- \( \rightarrow \) peksame, or the word-based inference of \( \text{peksame} \) from, say, \( \text{epeksa} \). Thus, in a realizational approach to inflection, the grammar of a language determines the sets of properties with which a base may be associated, and for each such property, the morphology of a language defines the word form realizing this association.

From the above, it is clear that lexical theories (both incremental and realizational) are syntactocentric in the sense that the structure of inflected word forms is defined by independently motivated principles of syntactic structure (e.g. phrase structure rules or head movement). Such morpheme-based approaches to inflectional morphology are governed by the idea that an inflected word form’s grammatical and semantic content is determined by its representation as a combination of morphemes. The usefulness of morphematic factorization in morphological analyses has led to the assumption that morphology is basically a system of syntagmatic relations among morphemes (below the word level) and, thus, paradigms are nothing more than an epiphenomenal effect on defining the morphology of a language. Consequently, in lexical theories, paradigms are perceived as being equivalent to lists of related phrases or sentences having no status as theoretical objects and playing no essential role in morphological analyses.

On the other hand, in frameworks of inferential theories (both incremental and realizational), morphology is not just a system of syntagmatic linear arrangement of minimal meaningful elements, but also includes systematic relations in the paradigmatic dimension. That is, an adequately sufficient account of the morphology of a language should not be limited to accounts for word-internal morphotactics, but also for a variety of systematic relations among distinct cells in the same paradigm and among cells in distinct paradigms as well. Hence, paradigms do play an important role in organizing related word forms or/and affixes. In such paradigm-based theories, the primary object of analysis in the field of inflectional morphology (not only for the linguistic analyses but also for the language acquisition) is the paradigm’s structure rather than merely the word’s structure. On that account, paradigms are perceived as
central to the definition of an inflectional system and not as the epiphenomenon that they are assumed to be in lexical theories of morphology.

In recent years, many linguists have arrived at the belief that inflectional paradigms play a significant role in the definition of a language's grammar and lexicon. Research and studies investigating inflectional paradigms’ significance and properties are directed in many linguistic fields, such as grammatical theory (e.g. Stump 2001; Blevins 2006; Ackerman et al. 2009; among others), language typology (e.g. Carstairs 1987; Baerman et al. 2010; Stump & Finkel 2013; among others), psycholinguistics (e.g. Orsolini et al. 1998; Sonnenstuhl et al. 2000; Dressler 2000; Clahsen et al. 2001a; Bittner et al. 2003; Baayen & Schreuder 2003; Bittner et al. 2003; Dabrowska 2004, 2005; Milin et al. 2009a,b; Milin et al. 2009a; Labelle & Morris 2011; among others), historical linguistics (e.g. Maiden et al. 2011; Fertig 2013; among others), and computational linguistics (Beesley & Karttunen 2003; Brown and Hippisley 2012; among others).

Our analysis involves a computational model which adapts the framework of a paradigm-based theory, in the sense that fully inflected forms are mutually related through possibly recursive paradigmatic structures, defining associative relations between forms (see also Burzio 2004). In this model, unlike systems where a set of rules generates or processes fully inflected forms based on a base form (or/and a set of inflectional affixes), the inflected word forms are somehow viewed as being realizations of a lexeme. This point of view enables both the investigation of relations among the forms of the lexemes and their grouping into sets of lexemes whose forms are in the same relations or, in other words, their grouping into paradigms.

2.3 The status of paradigms in paradigm-based approaches to inflectional morphology

The central role of paradigms in word formation, mainly within Word-and-Paradigm (WP) approach to morphological description, has been put forward since the fifties, when Hockett (1954: 210) mentioned it in passing, and Robins (1959) presented its leading ideas. The notion of paradigm assumed through WP (e.g. Robins 1959; Matthews 1965, 1972, 1991; Zwicky 1985; Anderson 1992; Aronoff 1994; Beard 1995; Stump 2001), promotes the whole word, contrasted with any other unit, to be the basic element of analysis and the minimal unit associated with meaning. Hence, the lexicon is assumed to include only whole words and not individual morphemes (Matthews 1991: 187). That is, speakers deal with words, store them as a whole and rearrange them as required, based on other words as models. The predominance of word leads to two theoretical notions of the term:

(i) the *lexeme* which refers to all morphosyntactic content with which a word may be associated;
(ii) the *grammatical word* representing the form realized by the combination of a lexeme and any set of morphosyntactic properties.
The prominent status of words in this approach is not just a methodological commitment lying on epistemological demands, classification purposes or theoretical analyses. Their status is rather related to their stability and solidity (compared to sub-word units) for investigating grammatical relations and their informativeness, in the sense that one inflected form could be predicted by another one (Robins 1959; Matthews 1991). Consequently, the primacy of words as essential units presupposes and concurrently licenses the assumption that paradigms are essential to the definition of the inflectional morphology of a given language. That is, the central role of paradigms follows the reliable implications provided within its essentially uniform and fixed dimension which infers (and is inferred by) the solid and usually convergent dimension of words within them (Blevins et al. to appear). In other words, the grammatical information associated with a whole word form does not only make easier the identification of its own substance, but also facilitates highly reliable inferences about the substance of other forms belonging to the same paradigm.

In this respect, the definition of inflectional morphology in terms of paradigms seems appealing, because many generalizations about inflection are actually generalizations about paradigms (Stump 2016: 8-23, 27-29). A paradigm reflects the formal and semantic relations among fully inflected word forms by representing a particular class of related forms of the same lexeme (i.e., conjugation or declension) and, thus, it is generalizable to all members of that particular class. At the same time, a paradigm for a given class of related forms of the same lexeme also participates in relations with other paradigms representing different classes of lexemes of the same grammatical category (Beecher 2004: 9-10). Following these lines, we may understand many phenomena exhibited by real morphological systems such as syncretism, deponency, overabundance and suppletion, occurring within paradigm cell(s) (see also Stump 2016). As a result, in WP approaches, the definition of the inflectional morphology of a language is actually equal to the definition of its inflectional paradigms.

The assumption that words and paradigms are essential to the definition of inflectional morphology does not mean that full sequences of independent word forms and, subsequently, full inflectional paradigms are stored in the lexicon. That is to say, inflectional paradigms are not expected to display properties of inflectional tables which exemplify ordering relations that are unrelated to the concept of paradigm structure (Stump 2016: 29). Actually, their central role reflects their predictiveness value in morphological organization. Specifically, highly interconnected networks presented within an inflectional system enable its factorization in terms of exemplary paradigms and sets of principal parts. Each known word form may be predictive for other unknown forms of the paradigm, yet principle parts ensure the predictability of particular forms related to it, thus reducing uncertainty. As a result, we can perceive that the way a paradigm is stored in the lexicon depends on the entropy, predictability, and predictiveness of the parts that a paradigm consists of (Stump & Finkel 2013: 295-313 and the references therein). Perceiving gradations of inter-predictability among forms of words is a differentiating factor between morpheme-based and word-based
approaches, especially in terms of their adequacy for explaining the acquisition and processing of complex morphological systems (Blevins et al. to appear).

Nevertheless, the terms of morpheme-based approaches (e.g. root, stem or affix) are still used as regards the relations between and among fully inflected word forms. Interestingly, modern WP approaches following mainly Matthews (1965, 1972) are not exclusively word-based (e.g. Extended WP model: Anderson 1982; A-morphous Morphology: Anderson 1992; Paradigm Function Morphology: Stump 2001; realization-based and lexeme-based approaches: Zwicky 1985; Aronoff 1994; Beard 1995). In these models, while words are still essential for grammatical meaning, their morphotactics are more aligned with morphematic approaches, in the sense that fully inflected word forms are composed by more basic constituents. Hence, recent WP approaches to inflectional morphology can be characterized as morphosyntactically word-based, but morphotactically constructive (Blevins 2006).

In this perception, following Blevins (2006) and Blevins et al. (to appear), the lexicon can be presented as a repository of word items analysed into sub-word units, and grammatical words are composed from these units though the application of realizational rules. Thus, paradigm cells represent (morphosyntactic) properties of words, but word forms realizing these properties are derived. Thus, internal structure is important because the arrangements of sub-word units reflect patterns where classes of words participate in. In fact, words themselves are organized as fundamental parts of these patterns and it is these set of emergent patterns that defines the organization of morphological systems.

Following these lines, paradigms are assumed to capture morphological relations in two dimensions: in the syntagmatic dimension, a complex word form in the paradigm of lexeme is related to a more basic stem (or root) of the lexeme through a series of operations introducing the inflectional exponents of a word form; in the paradigmatic dimension, a complex word form systematically contrasts with other complex word forms in the paradigm of a given lexeme. Likewise, morphological units are presumed to demonstrate two kinds of structure; the (linear) syntagmatic formal structure (the internal morphology) of single units; the (vertical) relational structure (the paradigmatic oppositions) of a network of units. For WP models, both dimensions are relevant in several ways and provide us with a framework to analyze lexical phenomena, such as inflection, and determine the organization of the mental lexicon (Ježek 2016: 162). Therefore, it is possible to represent the lexicon of a given language as a form-based (self-)organized network in which lexical structures are generated as a consequence of the paradigmatic relations among fully inflected word forms.

3. Paradigm structure of MG conjugation

Since the early 1960s, MG conjugation has been a central subject within various theoretical approaches. The two former theoretical descriptions of MG conjugation, by Hamp (1961) and Koutsoudas (1962), exemplify Item and Arrangement models.
Lexical self-organization in the acquisition and processing of MG conjugation: an artificial neural network approach

(morpheme-based morphology) for their morphological analysis. Later, Matthews (1967) provides an analysis within WP approach. Philippaki-Warburton (1970) attempts to incorporate the notion of syntactic features and to present the morphophonemics of MG in the framework of generative phonology. Following the German structuralist tradition (e.g. Seiler 1958), Babiniotis (1972) argues that a synchronic morphological analysis should take into consideration the latent forces of the language, which may trigger a restructuring of the paradigm. Again, Philippaki-Warburton (1973), in her subsequent analysis, proposes a compromise between a generative approach and a WP model. In more recent years, Janda & Joseph (1992) have proposed an analysis of Greek verbal forms in terms of morphological constellations of words and redundancy statements, such as meta-templates. Ralli (2005, 2007) analyses the Greek verbal system according to the lexical morphology model (Lieber 1980; Selkirk 1982; Kiparsky 1982; Mohanan 1986), following the idea that all morphological combinations occur within the Lexicon, which is not perceived as a simple repository of information, but as a dynamic component with word-formation rules and with lexical phonological rules responsible for word-internal phonological changes.

The below theoretical analysis follows the seminal work provided by Ralli. In particular, the following description is based on Ralli’s analysis of MG conjugation and issues relevant to this (for inflectional features: Ralli 1998, 1999; for word structure: 2001: 118-127, 2003: 77-85, 2005: 106-116, 123-138, 232-240, 258-263; for ICs and allomorphy: 2007: 123-137, 2013: 277-279) and Mackridge’s (1985: 164-170) description of phonological and non-systematic morphological stem alternation. As I have already stated in the introduction, I will deal only with active, voice indicative paradigms.

### 3.1 Stem alternation related to inflection

The MG language belongs typologically to the family of the so-called fusional languages. In this type of languages, full forms always require the combination of a stem and an affix. Stems can be either morphologically simple or complex and are used as a base to form other inflectionally-related word forms. Paradigms normally exhibit a high degree of systematic formal redundancy, both intra- and inter-paradigmatic. However, generally in MG, there are verbs that show form variation in their inflectional paradigms. The relation between formally varying stems of the same lexeme, or stem alternation, is a common fact in the inflection of a great many languages. Stem alternation can be purely morphological, purely phonological or unconditioned. The relation between formally stem alternants is the result of a variety of factors, as summarized on Table 1.
The MG verb is based on the feature of aspect and it seems that there are some peculiarities related to this, causing extended stem alternation.

Allomorphs are stem alternants that are in complementary distribution. In MG, allomorphy is phonologically or morphologically-conditioned. The first case refers to stem alternations that are interpreted as the result of application of phonological rules. Consider the inflected forms of a verb like ‘ɣrafo ‘I write’ in (1):

\[
(1) \quad \text{ɣrafo ‘I write’} \quad \sim \quad \text{ɟrap-s ‘I wrote’}
\]

In this example, the verb ‘ɣrafo ‘I write’ displays two stem variants, /ɣrafo/ and /ɟrap/, depending on the phoneme that follows its stem-final consonant. If this phoneme is the [+continuous] /s/ of the aspectual marker, a dissimilation rule transforms the [+continuous] /f/ into the [-continuous] /p/. Thus, /ɣrafo/ is the basic stem expressing the concept of ‘write’, and /ɟrap/ is the outcome of a phonological rule applied to it.

The perfective active stem of most verbs, except those with imperfective stem ending in a liquid, and some ending in a nasal, ends in -s-, which is responsible for the phonologically conditioned stem alternation. There are, in other words, verbs that have sigmatic (from the Greek letter σ = sigma) perfectives and others that have asigmatic perfectives. Verbs that have asigmatic perfectives have a perfective active stem identical to the imperfective (e.g. krin- ‘judge’, dieftlin- ‘direct’), except (in some verbs) for an alternation in the radical vowel (e.g. simen- ‘mean’ ~ siman-). Sigmatic perfectives may be considered as being formed by the addition of -s- to the stem-final vowel or underlying stem-final consonant of the imperfective.

Furthermore, there is a lot of intra-paradigmatic phonologically-conditioned stem alternation because of the endings starting with the frontal phonemes [i] or [e]. This context is sensitive to the phenomenon of palatalization, which is the single most important
phonetic phenomenon of the MG language (Newton 1976: 126-127). Hence, consonants, [k], [ŋ], [x] (and [g]) change from velar to the equivalent palatal one [c], [j], [c] (and [j]), if the vowel that follows is either [i] or [e] (e.g. ḏii ḳ-o ‘I manage’ – ḏii’c-is ‘you manage’, oḍi’y-o ‘I drive’ – oḍi’j-i ‘(s)he/it drive’, sime ‘tx-o ‘I participate’ – sime’teçete ‘you participate’).

In contrast, morphological allomorphs cannot be interpreted synchronically by the application of phonological rules and, therefore, these alternants are generally unpredictable. In this case, there is a further distinction between systematic and non-systematic allomorphy (Ralli 2007). For instance, in some verbs, there is a form difference between the stem that is used in the [+perfective] context and the stem used in the [+perfective] one. To illustrate this phenomenon:

(2) Systematic morphological allomorphs
   a. aya’p(a)-o ‘I love’ ~ a’yapi-s-a ‘I loved’
   b. fo’r(a)-o ‘I wear’ ~ fo’re-s-a ‘I wore’
   c. xa’l(a)-o ‘I demolish’ ~ xala-s-a ‘I demolished’

(3) Non-systematic morphological allomorphs
   a. ḏevy-o ‘I leave’ ~ e-fiy-a ‘I left’

A number of verbs displays the systematic stem variation shown in Example (2), according to which a X(a) stem variant is used in the context of imperfective forms, while a Xvowel stem variant appears in the context of perfective forms. Clearly, there is no synchronic phonological explanation for this form variation. The morphologically-conditioned allomorphy of this type requires a systematic pattern of perfective stem formation, namely X(a) ~ X + V (e.g. aɣap(a) > aɣapi-), where ‘X’ is a variable standing for the bare stem, ‘V’ stands for a vowel, and the subscripted ‘(a)’ indicates an optional ‘a’, forming a MG free variant of the imperfective stem (Ralli 2005, 2007). The variable V in the perfective stem can be instantiated as an i, e or a, and cannot be predicted synchronically from the bare stem.

Now consider a verb like the one in (3), where fevy- is in complementary distribution with fiy-. The former appears in the context of the [−perfective] aspectual value, while the latter is used in a context marked by [+perfective]. It should be noticed that fiy- does not result from the application of a rule, since it constitutes a different case from the variation displayed by verbal types like yraʃ- vs. yraʃ- and, furthermore, there is no any transparent systematic relation between the alternant, such as the case of X(a) ~ Xvowel allomorphs.

The case of verbs displaying non-systematic allomorphy (and the extreme case of suppletion) could be perceived as a defining feature of irregular verbs in MG. The irregularities appear not in the endings but in the formation of the stems, in the sense that the morphological allomorphs do not follow a systematic pattern. These verbs vary in their degree of irregularity, in that while a few have four different underlying stems, others have three, and others only two. Irregularities include deletion or addition of a sound or sounds, change of radical vowel, change of stem-final consonant, metathesis
of sounds, suppletion of one stem by another or combination of these irregularities within the paradigm of a single verb.

Allomorphs are morphemes which differ slightly regarding the sound form, sharing the same meaning (or function) and, as a result, the connection between them is more or less transparent. On the contrary, in the case of suppletion, there is no transparent connection between the morphemes, since the suppletive forms belong to different stems or/and lexemes. In fusional languages, the observation that suppletion is closely related to inflection (e.g. Dressler et al. 1987) is a commendable one. According to Ralli (2014), in MG conjugation, different stems could be used for the formation of the inflectional paradigms expressing the same verbal notion. Suppletion is found in many common verbs, in which the stem supplying the imperfective stem is completely different from that of the other stems. Examples are: le(ɣ)- ‘say’, perfective p- (i-p- in past and passive); vlep- ‘see’, perfective δ- (i-ð- in past and passive); and τro(y)- ‘eat’, perfective fag-.

The cases of stem alternation mentioned above underlie context constraints (see complementary distribution) and, thus, are considered as conditioned variants. However, some cells of MG conjugation are filled with doublets where the conditions of alternants cannot be defined. In the case of free variants, two different forms appear in the same cells. The cases where two different forms occur in the same morphological context are rare and usually explained by sociolinguistic criteria, such as the dialectal origin of the speaker. Furthermore, forms coexist in the same cell in times of language change. However, free variants are systematic in MG. More specifically, the cells of imperfect past of a particular verb class (see below) exhibit doublets systematically.

### 3.2 Inflectional affixes

As it is mentioned above, MG belongs to the fusional type of languages, in that the words of the major grammatical categories are complex formations containing stems and inflectional affixes which are portmanteau morphemes, combining more than one morphosyntactic feature. Inflection is generally realized as suffixation and, regarding verbs, at this process, the first constituent is always a verbal stem (simple or complex) and the second (or the final) is an (portmanteau) inflectional suffix, carrying the features of tense, person and number. Verbs are morphologically marked for voice, aspect, tense, person and inflection class, and are divided into two basic inflection classes. It is important to note that the features of aspect, voice and tense are overtly realized when they assume a specific value: overt forms exist for the perfective aspect, the passive voice and the past-tense. The values of [-perfective], [-passive], and [-past] are not related to specific forms, but are inferred by the co-occurrence of other features.

In MG, there are only inflectional suffixes and not prefixes. The augment e- in the past-tense of verbs is the only case where inflection could be considered to appear as a prefix, and in fact, it marked the past-tense in Ancient Greek. Nowadays, the use of e-is connected to the presence of stress on the grounds that it is absent in unstressed position. Since its occurrence is not compulsory in the past-tense, its inflectional status
is doubtful. In fact, in the literature, it has often been considered as a morphophonologically inserted element which is preposed to bi-syllabic verbal forms in the context of the past-tense (paradigms of imperfect and aorist). In particular, Babiniotis (1972), Kaisse (1982) and Ralli (1988) claim that the augment is nothing but a formative whose only function is to receive stress. More specifically, when the antepenultimate-syllable (third syllable from the end) stress law causes a left-hand stress shift outside the confines of the word, e- has the role of providing a cell to stress bi-syllabic verbal forms. For example, as shown in (4), e- is present since the combination of the stem and the inflectional ending accounts for two syllables.

(4) δέν-α > ῥδέν-α > ἐ-δέν-α > ἑδένα ‘I was tying’

### 3.3 The morphological organization of MG conjugation

The concept of ICs (referred to as declension or conjugation as well) has been recognized since ancient times as a very useful tool for the description and teaching of languages and, therefore, it was adopted as an important building unit of the lexicon by early Latin and Greek grammarians (see Matthews 1991). Inflectional systems are almost always organized into different inflection classes, usually exhibiting macroclasses or subclasses in cases where there is considerable overlap. These declensions and conjugations do not express morphosyntactic properties themselves, but determine how such properties are expressed. According to Wurzel (1984) and Carstairs-McCarthy (1994), relevant parts of a paradigm must be inspected to discover in which IC a lexical unit belongs. Consequently, since ICs are supported by paradigms, it must be assumed that several stems (or word forms) are listed in the lexicon someway.

As it has been mentioned, MG conjugation exhibits high degree of stem alternation both morphologically and phonologically-conditioned. Ralli (2005, 2007) proposed that morphologically-conditioned allomorphs which are not entirely phonologically dissimilar (as cases of suppletion) have a classificatory role, leading to the distinction of inflection classes and play an active role in paradigmatic organization in MG. Hence, Ralli (2007) suggests that the kind of morphologically-conditioned allomorphy exhibited by a paradigm functions as an *inflection-class demarcator* or like a schema (Bybee & Slobin 1982), since it determines the paradigmatic behavior of a class of verbs. Moreover, she points out the role of allomorphy as a significant factor to paradigmatic distinctness, as opposed to Carstairs (1987: 222-223) who claims that stem allomorphy is irrelevant to the identification of paradigms, to which only inflectional affixes should count.

More specifically, unlike nouns and adjectives whose IC is not predictable but an inherent lexical feature, the IC of Greek verbs is provided by the presence or absence of systematic stem allomorphy. Verbs are conjugated according to two inflection classes, the distinction of which is based on a stem-allomorphy pattern $X_{(a)} \sim X + V$. Verbs that do not adapt to the particular allomorphy pattern are predicted to inflect
differently from verbs that have it and it defines verbs belonging to IC1. There is also one more subdivision in IC2 depending on the presence or absence of allomorph $X_a$:

\[(5)\]

\[
\begin{array}{c|c}
\text{IC2} & \\
\hline
\text{IC2}_a & \text{IC2}_b \\
\hline
X_a & X_1 \\
\hline
\end{array}
\]

\[\text{aya}'p(a)-i \quad \delta i e' r-i \]

\[\text{‘(s)he/it loves’} \quad \text{‘(s)he/it divides’} \]

The main difference between the two subsets is that IC2$_b$ lacks the form -ɣ-$ for the marking of [-perfective] and uses only the inflectional suffix -us- (see below). However, IC2$_b$ is not so productive and its verbs are fewer than the verbs of IC2$_a$.

Paradigms classified into IC1 show allomorphy, but it is either unsystematic or phonologically-conditioned. Specifically, IC1 presents the following subcategories:

(i) absence of morphologically-conditioned allomorphs,
   e.g. ‘yráf-o ‘I write’ ~ ‘e-grap-s-a ‘I wrote’;

(ii) there is allomorphy but not systematic,
   e.g. ‘fevɣ-o ‘I leave’ ~ ‘e-fiy-a ‘I left’; and

(iii) there is suppletion,
   e.g. ‘tro-o - ‘I eat’ ~ ‘e-fay-a ‘I ate’.

Hence, it is useful to illustrate the cases of morphologically-conditioned stem alternation which is perceived to interact with the organization of verb paradigms in MG. For the illustration (Table 2), I base my analysis on Ralli’s classification (2005: 138) and Karasimos’ (2011: 75) description of Greek verbal allomorphy.

<table>
<thead>
<tr>
<th>Categories</th>
<th>-perfective</th>
<th>+perfective</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Present/Future</td>
<td>Imperfect (Progressive)</td>
</tr>
<tr>
<td>Category 1 (IC1)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>e.g. 'ði'ok-o ‘I prosecute’ ~ 'ðiok-s-a ‘I prosecuted’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category 2a (IC2a)</td>
<td>X$_{(a)}$</td>
<td>X$_{(a)}$</td>
</tr>
<tr>
<td>e.g. aya'p(a)-o ‘I love’ ~ a'ya-pi-s-a ‘I loved’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category 2b (IC2b)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>e.g. δié r-o ‘I divide’ ~ δié ere-s-a ‘I divided’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Category 3 (IC1)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>e.g. ‘fevɣ-o ‘I go away’ ~ ‘e-fiy-a ‘I went away’</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Categories</td>
<td>-perfective</td>
<td>+perfective</td>
</tr>
<tr>
<td>----------------------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td></td>
<td>Present/Future Progressive</td>
<td>Imperfect (Progressive)</td>
</tr>
<tr>
<td>Category 4 (IC1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e.g. <code>pern-o ‘I take’ ~ </code>pir-a ‘I took’ (tha) `par-o ‘I will take’</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Category 5 (IC1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e.g. <code>tro-o ‘I eat’ – </code>e-fag-a ‘I ate’</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Category 6 (IC1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>e.g. <code>le-o ‘I say’ – </code>ip-a ‘I said’ ~ (tha) p-o ‘I will say’</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

**Table 2:** Distribution of morphologically-conditioned stem change in active voice, indicative

Finally, assuming that the general structural pattern for the verb types is [Stem-(Aspect)-Tense/Person/Number] (Koutsoudas 1962; Ralli 2005), the paradigms of active present, imperfect and aorist indicative are presented by the following tables which illustrate the two inflection classes, with respect to the examples `dēno ‘I tie’ and ayá p(a)o ‘I love’. For clarity reasons, hyphens separate the verbal base from the augment and the inflectional suffixes.

**Active Voice**

<table>
<thead>
<tr>
<th>Present/Future Progressive</th>
<th>Imperfect (Progressive)</th>
<th>Aorist (Simple Past)</th>
<th>Simple Future</th>
</tr>
</thead>
<tbody>
<tr>
<td>`dēn-o</td>
<td>`e-dēn-a</td>
<td>`e-dē-s-a</td>
<td>`dē-s-o</td>
</tr>
<tr>
<td>`dēn-is</td>
<td>`e-dēn-es</td>
<td>`e-dē-s-es</td>
<td>`dē-s-is</td>
</tr>
<tr>
<td>`dēn-i</td>
<td>`e-dēn-e</td>
<td>`e-dē-s-e</td>
<td>`dē-s-i</td>
</tr>
<tr>
<td>`dēn-ume</td>
<td>`e-dē-ame</td>
<td>`e-dē-s-ame</td>
<td>`dē-s-ume</td>
</tr>
<tr>
<td>`dēn-ete</td>
<td>`e-dē-an</td>
<td>`e-dē-s-an</td>
<td>`dē-s-ete</td>
</tr>
<tr>
<td>`dēn-un(e)</td>
<td></td>
<td></td>
<td>`dē-s-un</td>
</tr>
</tbody>
</table>

**Table 3:** IC1 (stem paradigm: `dēn ~ ðe ‘tie’)

<table>
<thead>
<tr>
<th>Present/Future Progressive</th>
<th>Imperfect (Progressive)</th>
<th>Aorist (Simple Past)</th>
<th>Simple Future</th>
</tr>
</thead>
<tbody>
<tr>
<td>ayá p(a)-o</td>
<td>a ɣapa-y-a/ayá p-us-a</td>
<td>a ɣapi-s-a</td>
<td>ayá pi-s-o</td>
</tr>
<tr>
<td>ayá pa-s</td>
<td>a ɣapa-j-es/ayá p-us-es</td>
<td>a ɣapi-s-es</td>
<td>ayá pi-s-is</td>
</tr>
<tr>
<td>ayá pa-(i)</td>
<td>a ɣapa-j-e/ayá p-us-e</td>
<td>a ɣapi-s-e</td>
<td>ayá pi-s-i</td>
</tr>
<tr>
<td>ayá pa-me/ayá p-ume</td>
<td>a ɣapa-y-ame/ayá p-us-ame</td>
<td>ayá pi-s-ame</td>
<td>ayá pi-s-ume</td>
</tr>
<tr>
<td>ayá pa-te</td>
<td>a ɣapa-y-ate/ayá p-us-ate</td>
<td>ayá pi-s-ate</td>
<td>ayá pi-s-ete</td>
</tr>
<tr>
<td>ayá pa-ne/ayá p-un(e)</td>
<td>a ɣapa-y-an/ayá p-us-an(e)</td>
<td>ayá pi-s-an/</td>
<td>ayá pi-s-un(e)</td>
</tr>
</tbody>
</table>

**Table 4:** IC2ₐ (stem paradigm ayap(a) ~ ayapi ‘love’)

Lexical self-organization in the acquisition and processing of MG conjugation: an artificial neural network approach
### Table 5: IC2b (stem paradigm ðier ~ ðiere ‘divide’)

<table>
<thead>
<tr>
<th>Present/Future Progressive</th>
<th>Imperfect (Progressive)</th>
<th>Aorist (Simple Past)</th>
<th>Simple Future</th>
</tr>
</thead>
<tbody>
<tr>
<td>ðier-o</td>
<td>ðier-us-a</td>
<td>ðier-ere-s-a</td>
<td>ðier-re-s-o</td>
</tr>
<tr>
<td>ðier-is</td>
<td>ðier-us-es</td>
<td>ðier-ere-s-es</td>
<td>ðier-re-s-is</td>
</tr>
<tr>
<td>ðier-i</td>
<td>ðier-us-e</td>
<td>ðier-ere-s-e</td>
<td>ðier-re-s-i</td>
</tr>
<tr>
<td>ðier-ume</td>
<td>ðier-us-ame</td>
<td>ðier-re-s-ame</td>
<td>ðier-re-s-ume</td>
</tr>
<tr>
<td>ðier-ite</td>
<td>ðier-us-ate</td>
<td>ðier-ere-s-ate</td>
<td>ðier-re-s-ete</td>
</tr>
<tr>
<td>ðier-un(e)</td>
<td>ðier-us-an(e)</td>
<td>ðier-ere-s-an/ðier-re-s-an(e)</td>
<td>ðier-re-s-un(e)</td>
</tr>
</tbody>
</table>

### 3.4 Morphological regularity in MG conjugation

As I said above, past-tense in MG is a morphologically complex form, as tense interacts with aspect (perfective/imperfective). Consequently, MG makes a distinction between perfective and imperfective verb stems (Holton et al. 1997). The latter is formed with the imperfective stem itself (e.g. ‘yraf-o ‘I write’ – ‘e-yraf-a ‘I was writing’) and thus, it is morphologically simpler than the perfective one. The perfective stem can be either sigmatic, which involves a segmentable affix -s- or non sigmatic. Notably, verbs that take sigmatic forms are considerably more common in the MG language than those that take non-sigmatic forms (Stavtrakaki & Clahsen 2009a).

Ralli (1988, 2005) provided a classification of verb paradigms which is based on two criteria; firstly, the presence or absence of the sigmatic affix and, secondly, the presence or absence of (systematic) stem allomorphy. As a result, the below categories can be defined (Tsapkini et al. 2001: 282-283; Tsapkini et al. 2002a: 268; Tsapkini et al. 2002b: 318-319; Tsapkini et al. 2002c: 106-107; Tsapkini 2004: 613):

(i) an affix-based paradigm with the presence of aspectual marker -s-, which includes verbs with a predictable phonological stem change but not a morphological one (e.g. ‘lin-o ‘I untie’ ~ ‘e-li-s-a ‘I untied’, ‘yraf-o ‘I write’ ~ ‘e-yrap-s-a ‘I wrote’);

(ii) a mixed paradigm where active perfective past-tense forms are produced by applying the (rule-based) aspectual marker -s- to a (stored) systematic morphological stem alternant (e.g. ‘mi’ll(a)-o ‘I speak’ ~ ‘mili-s-a ‘I spoke’);

(iii) a stored paradigm including non-systematic stem alternation (or suppletive forms) or no stem alternation and lacking the aspectual marker (e.g. ‘pern-o ‘I take’ ~ ‘pir-a ‘I took’, ‘tro-o ‘I eat’ ~ ‘e-fag-a ‘I ate’, ‘krin-o ‘I judge’ ~ ‘e-krin-a ‘I judged’).

Consequently, the presence of the affix -s- is related to stems that show phonologically predictable and morphologically systematic changes*. By contrast, asigmatic past-tense

---

* The difference between predictable and systematic is that the former results from rules with no exceptions or differentiation and the latter is the outcome of a pattern whose result is not predictable but systematic. In particular, phonological stem alternants, on the one hand, are fully predictable, since they
forms are less systematic and non-predictable and include suppletive stem-changes as well, even though there are some tendencies and patterns.

Following this classification, we can perceive the (ir)regularity connected to verb conjugation as a cline. On the one hand, asigmatic forms can be perceived as irregular, including cases of (stored) unpredictable morphologically-conditioned stem alternants (either allomorphs or suppletive forms) or asigmatic past-tense forms which do not exhibit stem alternation at all. On the other hand, sigmatic past-tense forms are regular, since they involve a segmentable (rule-based) affix combined with a predictable phonologically-conditioned stem alternant. However, there is also an in-between category where the (rule-based) aspectual marker is combined with a (stored) non-predictable morphologically-conditioned stem allomorph. As a result, we could say that the gradient levels of regularity are the following (they are also summarized in Table 6 below):

(i) **regular cases**: sigmatic forms presented predictable phonologically-conditioned stem alternation (e.g. ‘li-n-o ‘I untie’ ~ ‘e-li-s-a ‘I untied’, ‘yra-f-o ‘I write’ ~ ‘e-yrwap-s-a ‘I wrote’);
(ii) **(semi-)regular cases**: sigmatic forms presented systematic stem alternation, following Ralli’s pattern X(a) ~ X_vowel (e.g. ‘mi-li(a)-o ‘I speak’ ~ ‘mili-s-a ‘I spoke’);
(iii) **irregular cases**: asigmatic forms presented non-systematic stem alternation, usually stem-internal change (e.g. ‘per-n-o ‘I take’ ~ ‘pir-a ‘I took’), asigmatic forms presented no stem alternation at all (e.g. ‘krin-o ‘I judge’ ~ ‘e-krin-a ‘I judged’) and inflected forms presented suppletion (e.g. ‘tro-o ‘I eat’ ~ ‘e-fag-a ‘I ate’).

<table>
<thead>
<tr>
<th>Stem Formation</th>
<th>Aspectual Marker</th>
<th>Relations</th>
<th>(Ir)regularity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purely Phonological</td>
<td>+</td>
<td>Phonological Allomorphs</td>
<td>Regular</td>
</tr>
<tr>
<td>(predictable)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Purely Morphological</td>
<td>+</td>
<td>Systematic Morphological Allomorphs</td>
<td>(semi-)Regular</td>
</tr>
<tr>
<td>(non-predictable)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Non-Systematic Morphological Allomorphs</td>
<td>Irregular</td>
</tr>
<tr>
<td>Null</td>
<td></td>
<td>Suppletive Forms</td>
<td></td>
</tr>
</tbody>
</table>

*Table 6: Stem alternation and (ir)regularity*

In conclusion, in Greek verb inflection, regularity is based on the important distinction between sigmatic and non-sigmatic perfective stem forms; the former group of verbs are the outcome of phonological rules of Greek and systematic morphological stem alternants are of this kind because they follow the pattern X(a) ~ X + V, where the vowel could be i, e or a. Thus, they are systematic, but not predictable because the vowel of the resultant stem allomorph cannot be predicted given the bare stem.
contains an -s- perfective affix and the latter group lacks -s-. In this dissertation, we consider sigmatic forms as a case of regular inflection, while asigmatic verb forms as typical cases of irregular inflection. Sigmatic verbs are considered to be regulars because their aspectual marker is combined with either phonologically predictable or morphologically systematic stem-changes; non-sigmatic past-tense forms, by contrast, exhibit properties of irregular inflection, in that they involve unpredictable stem-changes and no segmentable (perfective past-tense) affix. Finally, the sigmatic perfective past-tense is more frequent and, due to the -s- suffix and the systematic stem allomorphy, more transparent than asigmatic verb forms which do not involve a segmentable perfective past-tense affix and are more idiosyncratic.

4. Word learning and processing: psycho-linguistic approaches

The issue of regularity in inflectional morphology is mainly represented by the on-going debate between dual-route and single-route models in the field of morphological learning and processing. In general, the dual-route models claim that regular words are processed on-line (by rules), while morphologically irregular words are stored as wholes in the lexicon. On the other hand, single-route approaches hold that both morphologically regular and morphologically irregular words are processed by a single mechanism. These assumptions have sparked a great amount of studies and research prolonging this debate.


4.1 Single- vs. dual-route models

Inflectional morphology is the basic domain upon which the debate between single- and dual-route approaches is developed, because it typically exhibits the dissociation between regularity and irregularity, providing a central testing ground for contradicting views of language. Specifically, inflectional morphology appears to play an important role in shaping the neural responses to the processing of regularly and irregularly inflected forms (Tyler et al. 2002; Shapiro et al. 2006). Moreover, past-tense acquisition
in English is characterized by a U-shaped pattern of development dependent on the above disassociation (Plunkett & Marchman 1991). The assumptions that neural representations of grammatical categories may be mediated by this morphological distinction and the developmental trend of past-tense in English provided the chance of distinguishing between regular and irregular morphological forms and the psychological mechanisms associated with them. As a result, the issue of whether regular and irregular morphology can be processed and derived by a single or a dual mechanism has led to one of the longest and most dynamic debates within linguistic research, where several approaches were proposed in order to explain this dissociation.

In generative school, Chomsky and Halle (1968) proposed that both regular and irregular past-tense forms are governed by a set of rules and minor rules, respectively. However, it was relatively questionable that the derivation of past-tense forms occurs by the application of rules exclusively (Bybee & Moder 1983). In this context, Rumelhart and McClelland (1986) proposed that both regular and irregular forms could be learned and derived by a connectionist model, a single computational mechanism which was not fully rule-based and assembles all the inflections of a language. Here, the rules were replaced by weighted connections in a connectionist pattern associator, where the acquisition is data-driven on the basis of the connection weights. This network model recognizes patterns between input and output forms and, thus, the role of morphology is captured by a general processing system that both irregular and regular forms are treated in the same manner; in a way that morphology is not explicitly represented but emerges from the correspondence between form and meaning. As a consequence, the main position of this approach is the total absence of the regular-irregular distinction in the processing system (Rumelhart & McClelland 1986; Plaut & Gonnerman 2000).

A critical position on the connectionism by Pinker and Prince (1988) has paved the way for the approaches which posit two different types of processes to word processing and are known as dual-route models (Clahsen 1999; Pinker 1991; Prasada and Pinker 1993; Pinker and Ullman 2002; among others). The dual-route approaches maintain the rules as the basis for regular forms, while trying to provide a more sufficient account for irregulars. The general idea is that regular words are always processed in a decompositional way, whereas irregular forms are stored as wholes (Pinker & Prince 1994). As a result, irregular and regular forms are subserved by different architectural components: the one route is governed by rules and enables the formation of regular forms and the second is related to a memory system of irregular forms, those that entail an idiosyncratic, non-systematic (usually stem-internal) alternation. This distinction is claimed to have certain consequences for the way these forms are processed. In particular, the second route is taken if and only if the first fails to find any matching access entry in the lexicon.

As with other debates in the field of psycholinguistics and language acquisition, behavioural evidence is not clear in the sense that both competing theoretical approaches could account for the processing of regular and irregular morphology. For instance, both single- and dual-route theories, explain the U-shaped pattern in language acquisition of English past-tense. As theoretical scenarios become more complex, it is
useful to examine their descriptive and interpretative adequacy in languages with richer inflectional system than English, providing a more complex ground for the contradicting approaches. Evidence for and against both models has been provided investigating morphological processing in other languages, such as German (Bybee 1991; Clahsen et al. 1992, 2001b; Bybee & Newman 1994; Marcus 1995; Nakisa & Hahn 1996; Penke et al. 1997; Hahn & Nakisa 2000; among others), Italian (Burani et al. 1984; Caramazza et al. 1988; Orsolini et al. 1998; Say & Clahsen 2001; Eddington 2002; among others) and Spanish (Rodriguez-Fornells et al. 2002; Clahsen et al. 2002; Yaden 2003; Brovetto & Ullman 2005; Eddington 2004; among others). In the next chapter, I will focus on literature pertaining to MG morphological regularity and the mechanisms suggested to be associated with.

4.2 Morphological (ir)regularity: experimental evidence from MG conjugation

For many years, issues of morphological (ir)regularity have traditionally been investigated through the prism of morphological competence issues, with particular emphasis on aspects of the internal structure of complex words (Bloomfield 1933; Bloch 1947; Chomsky & Halle 1968; Lieber 1980; Selkirk 1984; among others). In this framework, one of the strongest positions is that morphologically, phonologically or/and semantically transparent words are always processed through their constituent elements, whereas irregular idiosyncratic forms are stored as wholes (Pinker & Prince 1994). The definition of morphological regularity is deeply related to this general idea and, in this line, there are several psycholinguistic studies focusing on this morphological distinction investigating regularity of Greek verb forms.

The definition of regularity following the peculiar characteristics of word-based languages, such as English, is problematic in the case of stem-based languages, like MG. On the basis of Ullman and colleagues’ (1997) assumption that the past-tense formation of regular verbs in English requires on-line application of an affixation rule (e.g. talk > talk-ed), while irregular past-tense forms, involving stem allomorphy (e.g. drink > drank), are retrieved from the lexicon, Terzi et al. (2005) investigated the production of regular and irregular verbs in the past-tense by Greek-speaking patients with Parkinson’s Disease. In contrast with Ullman et al.’s (1997) prediction of a better performance with the formation of the past-tense of irregular verbs, they found that the behavior of patients with Parkinson’s Disease is not different at the formation of regular and irregular past-tense. According to them, this is not a surprising finding, given the properties of the Greek past-tense morphology, where the regular/irregular distinction, in Ullman and colleagues’ sense, is not valid. The reason is that there are not forms which are exclusively based on the application of grammatical rules and on the retrieval of a stored form from the lexicon (Terzi et al. 2005: 298). Instead, since all words are inflected, the decomposition route should apply for both regulars and irregulars at least in the initial stages of lexical access and for the inflectional suffixes of tense and agreement that are common to both (i.e., not the rule-based aspactual marker -s). As a
result, MG does not show the clear-cut distinction that English exhibits in the process of forming the past-tense of regular vs. irregular verbs.

Hence, the difference between regular and irregular (rule-based and memory-based) morphological processing in MG is what happens at next stages, after the detachment of the common inflectional affixes of tense and agreement, e.g. what happens at the level of stem. This is significant for a stem-based language, such as MG (Tsapkini et al. 2002b: 317). As a result, psycho-linguistic research has focused on the important distinction between sigmatic and asigmatic verb forms. In this framework, there are several studies investigating different populations by employing different methods. There are studies of individuals with Specific Language Impairment (Kehayia 1997; Stamouli 2000; Mastropavlou 2006; Smith 2008; Stavrokaki & Clahsen 2012), focal brain lesions (Konstantinopoulou et al. 2013), Williams Syndrome (Varlokosta et al. 2008; Stavrokaki & Clahsen 2009b) and Down Syndrome (Stathopoulou & Clahsen 2010), typically developing children (Stavrokaki & Clahsen 2009a) and aphasic/agrammatic patients (Kehayia 1988; Tsapkini et al. 2001, 2002a,b). All of them suggest a two-way distinction (between sigmatic and asigmatic past-tense forms) within Greek inflectional system which is consistent with the broad distinction between combinatorial and memory-based morphological processes posited in dual-mechanism morphology; with the former being involved in -s- suffixation of sigmatic perfective past-tense forms and the latter in the unpredictable allomorphic stems of asigmatic perfective past-tense forms. However, a careful review of the Greek data questions this sharp distinction. In particular, the Greek data also provide the case of a mixed paradigm where both non-predictable stem alternation and affixation are simultaneously present in the formation of past forms (e.g. miˈl(a)-o ‘I talk’ ~ mili-s-a ‘I talked’, ðin-o ‘I give’ ~ ðo-ðo-s-a ‘I gave’). This advocates for Tsapkini et al.’s (2002c: 104) claim that both the role of stem allomorphy and the presence of affixes are important for the definition of regularity for languages with rich inflectional system, like MG.

Psycholinguistic evidence provides robust empirical support to the assumption that the human lexical processor is sensitive to the subtle differences as regards the classification of the Greek verb system. Morphological regularity should not be considered as an epiphenomenon of the design of the human language ability or the dualism between rule-based and memory-based routes. Rather, it is more efficient to be perceived as the graded result of the varying interaction of morpho-semantic, morphotactic, and phonological factors. More specifically, the lack of full formal nesting between imperfective and perfective stems (as in ðuˈlev-o ‘I work’ ~ ðulep-s-a ‘I worked’) appears to have an extra processing cost for speakers. Tsapkini and colleagues (2002c) compared priming evidence obtained from Greek past-tense forms matched for orthographic overlap (50%) with present forms, in two different conditions of Stimulus-Onset Asynchrony (SOA): a short SOA (35 ms) and a long one (150 ms). In the initial stages of lexical access, facilitation is elicited for both irregular and regular verb primes (compared to unrelated primes). However, not all forms seem to benefit from the augmented processing time permitted by longer SOA: non fully-nested allomorphic stems (e.g. ðuˈlev-o ‘I work’ ~ ðulep-s-a ‘I worked’, or ðern-o ‘I take’
~ ‘pir-a ‘I took’) elicit significantly less priming than do fully-nested allomorphs (e.g. mi’l(a)-o ‘I speak’ ~ ‘mili-s-a ‘I spoke’), with intermediate cases of formal overlap (e.g. ’lin-o ‘I untie’ > ’e-li-s-a ‘I untied’) getting intermediate facilitation effects. This graded behaviour is confirmed by Voga and Grainger (2004), who found greater priming effects for inflectionally-related word pairs with larger onset overlap. Along the same lines, Orfanidou and colleagues (2011) investigated the effects of formal transparency/opacity and semantic relatedness in both masked (short SOA) and delayed (long SOA) priming for derivationally-related Greek word pairs. They report evidence that semantically transparent primes produced more facilitation in delayed priming than semantically opaque primes, while orthographically transparent primes produced more facilitation in masked priming than orthographically opaque primes.

Consequently, the Greek data and the experimental results support the idea of “degrees or gradations of morphological regularity, as a connectionist single-mechanism and analogy-based approach would predict, rather than qualitatively different operations based on the distinction between regulars and irregulars, as the dual-mechanism approach would predict” (Tsapkini et al. 2004: 612). This approach surpasses the issues occurring in defining morphological regularity in terms of stem-change and affixation, where irregular forms involve change at the stem, whilst regular forms involve no stem-change, but rather an affix addition (Clahsen 1999; Say & Clahsen 2001). “This asymmetry in the definition of regularity in poorly inflected languages, e.g., English, does not allow to determine whether the presence of stem-change, i.e., stem allomorphy, or the presence of affixes is the crucial characteristic of morphological regularity and the one that affects processing.” (Tsapkini et al. 2002c: 104). Regarding this, the MG language provides the opportunity to take some distance from these dichotomies (rule-based vs. stored-allomorph mechanisms) not only because all verb forms involve affixation and stem-change but also because of the presence of a category of verbs that combines both of these mechanisms (stored-allomorphy + rule-based suffixation).

Summarizing, under these accounts, the Greek verb system is described as a two-way system where the regularity of verb forms depends upon the presence or absence of the perfective past-tense affix -s-. However, MG presents an interesting new dimension because regularity is not an all-or-nothing concept in the language, since past-tense formation yields a more complex picture in MG than other languages, such as English (Tsapkini et al. 2004: 616). So, the MG language provides the chance for a revision of the traditional dichotomy between grammar and lexicon, process and storage through a single mechanism. However, even if the clear-cut distinction proposed by the dual-route models is rejected, from the connectionist perspective, the original Rumelhart and McClelland model is no more at issue, but rather the single mechanism philosophy behind it. Thus, it is significantly important to determine the properties of such a mechanism and the evidence supporting a view of the lexicon as an emergent integrative system (Corbett & Fraser 1993; Wunderlich 1996; Dressler et al. 2006).
4.3 Storage vs. processing

In this dissertation, I follow a paradigm-based approach to morphological competence, according to which the acquisition of the inflectional system of a language is relevant to a number of constraints about the way that paradigm cells should be filled (see also ch.2). These constraints on both word acquisition and processing are the emergent result of the relational structure of all forms constituting a paradigm (Burzio 2004). In line with this view, there are psycholinguistic and neuro-physiological evidence supporting a view of the mental lexicon as a parallel distributed processor which operates on the basis of a dynamic storage, in which emergent redundant relations among multiple bases of fully stored forms play a significant role for both morphological processing and acquisition (McClelland & Elman 1986; Norris 1994; Bybee 1995, 2002; Hickok & Poeppel 2004; Shalom & Poeppel 2008; among others).

In the recent psycholinguistic literature on morphological competence, the mental lexicon is observed as a dynamic repository, in which fully stored word forms constitute the minimum units for morphological processing and the morphological acquisition/processing emerges from the relations among the morphologically-related forms (Bybee 1995, 2002; Burzio 2004; among others). The basic idea is that the organization and structure of the morphological lexicon should involve redundantly specified knowledge in order to maximize the efficiency of lexical access and retrieval, using both general and specific information at the same time (Post et al. 2008; Libben 2010). Thus, it is perceived as a dynamic memory system (Elman 2004), a parallel network of fully memorized word forms (Bybee 1995), whose boundaries between processing and storage are blurred and overlapped, being actually two sides of the same coin (Marzi & Pirrelli 2015: 497). This view of the mental lexicon is directly connected to paradigm-based approaches, in the sense that fully inflected forms are redundantly stored and mutually related through possibly recursive paradigmatic structures, defining entailment lexical relations between forms (Bybee 1995; Pirrelli 2000; Burzio 2004; Blevins 2006; Baayen 2007). There is no distinction between regular and irregular inflected forms (Baayen 2007; Marzi 2014). In fact, the way that words are represented in the lexicon reflects the way that they are dynamically processed, accessed and retrieved, “with levels of entrenchment and levels of resting activation being a function of the probabilistic support words receive from repeatedly successful processing steps” (Marzi et al. 2016: 82). As a result, morphological regularities are the outcome of independent principles of lexical organization and, therefore, the mental lexicon is perceived as an emergent integrative system (Corbett & Fraser 1993; Wunderlich 1996; Dressler et al. 2006).

The dynamicity of mental processes characterizing lexical competence (Bybee 1995) is also supported by neuro-linguistic data which promote a blurred distribution between processing and storage. In particular, considering that the implicated brain areas for language processing and working memory are significantly overlapped and that the latter could be perceived as the temporary activation of long-term memory structures (Hickok & Poeppel 2004; Shalom & Poeppel 2008), it could be supported that the responsible structures for the word processing are also responsible for its
representation in the long-term memory (Marzi & Pirrelli 2015: 497). As a result, the dual-route model of the grammar-processing relation is rejected, since the primacy of processing (grammar) over storage (lexicon) is based on a traditional view of lexical storage as costlier than computational operations (Baayen 2007). This view is questioned by alternative models which claim that lexical knowledge is organized in such a way to maximize the processing efficiency and opportunities regardless of storage constraints (Libben 2005). For instance, fully stored forms in the lexicon could respond better to some tasks, such as lexical decision, whereas lexical decomposition might be used efficiently for other tasks, like lexical acquisition. The best solution for a complex biological system, like the brain, is to entertain both organizational strategies simultaneously and make them compete on different tasks (Marzi and Pirrelli 2015: 495).

In this perspective, the way words are memorized could reveal the cognitive mechanisms and principles that determine word processing and lexical organization. In particular, the investigation of processing functions and low-level memory would not only evaluate their involvement in morphological processing but also account for it. From this perspective, the emergence of the morphological structure in the mental lexicon could be observed through the analysis of the dynamic interaction between lexical representations and distribution capacity, and perceptual discrimination and degrees of regularity in input data (Marzi & Pirrelli 2015: 496). For this purpose, artificial neural networks and self-organising maps are promising, because they provide experimental ways of testing this interaction under different input conditions and parameters (Pirrelli 2007).

5. A neuro-computational approach to MG conjugation

Until now we have seen theoretical linguists’ investigation about the kind of linguistic representations and features speakers must acquire so as to master the inflectional morphology of a language (ch.2 and 3), psycho-linguists’ examination of these findings via the language performance and their neuro-cognitive footprint (ch.4) and now we will see how these results could come about algorithmically via a neuro-computational implementation of a detailed model simulating the interaction between linguistic knowledge and psycho-cognitive and neuro-functional constraints.

A computer simulation is a mathematical model of a phenomenon represented in the form of an information processing computer program. Of course, a simulation of a specific brain function is not a brain. Thus, the presence of information processing does not imply a relationship between the program and the brain (Searle 1980: 6). However, if the output of a computer model of a specific brain function was connected to a machine which could mimic or perform all the cognitive functions that we intuitively associate with human minds, it would be reasonable to suggest that the system, taken as a whole, represents something closer to a real brain. It could be described as an artificial brain. That is to say, connectionist models manipulate electro-numeric patterns in a manner which is substantially closer to the electro-chemical pattern manipulation
of the brain than conventional computer programs, which are in turn closer than mechanical automata (Dyer 1990: 8). Thus, networks may support many complex systems and their theoretical and computational analysis contributes to highlighting issues related to their organizational and functional dynamics. Furthermore, artificial neural networks can be used to model many various real systems, where the units of the system are nodes, and interactions between units can be modelled as adaptive processes resulting to emergent behaviour (Marzi & Pirrelli 2015: 498).

From this angle, there is an ongoing view that computer models can significantly help not only for a clearer description but also for a better evaluation of the functional models of the mental lexicon, because they allow us to test the behaviour of a complex system under different parameters and input conditions in many experimental ways (Pirrelli 2007). Thus, artificial neural networks and, particularly, self-organising connectionist models, are not limited in hypothesis testing, but they can also clarify “the emergence of complex lexical representations from highly-interconnected relations taking place at the word level across different time scales, providing the intermediate level of scientific inquiry that bridges the gap between low-level, interactive brain processes and high-level language knowledge and language behaviour” (Ferro et al. 2010; Ferro et al. 2011; Marzi et al. 2012; Pirrelli et al. 2014; Marzi et al. 2014; Marzi et al. 2015).

However, despite the fact that there is a great interest and a large body of theoretical literature on computational models incorporating such a complex interaction between storage and processing (e.g. McClelland & Rumelhart 1981; Pollack 1990; Plaut et al. 1996; Harm & Seidenberg 1999; Coltheart et al. 2001; Botvinick & Plaut 2006; Perry et al. 2007; Sibley et al. 2008), there are only a few premising these properties (e.g. Li et al. 2007; Mayor & Plunkett 2010; Althaus & Mareschal 2013). Oddly, classical connectionist networks have failed to provide an alternative view of the interaction between grammar and lexicon. The latter seems to have no place in the classical computational frameworks of connectionism, which are attached to the basic rule-based approach for morphological inflection, thus reflecting a neurally-inspired perception of inflection rules (see McClelland & Patterson 2002; Ferro et al. 2010: 209-212; Pirrelli et al. 2011: 139-141; for a review).

Temporal Self-Organising Maps (Ferro et al. 2010; Ferro et al. 2011; Marzi et al. 2012; Pirrelli et al. 2014; Marzi et al. 2014; Marzi et al. 2015; Marzi & Pirrelli 2015; Marzi et al. 2016) are a neuro-biologically inspired computational model of lexical memories, in the sense that it simulates “the behaviour of brain maps, medium to small aggregations of neurons in the cortical area of the brain, involved in selectively processing homogeneous classes of data” (Pirrelli et al. 2011: 142). Therefore, they define an interesting class of general-purpose memory models, exhibiting a non-trivial interaction between short-term and long-term memory processes. At the same time, they simulate incremental processes of topological self-organization whereby lexical sequences are arranged in maximally predictive hierarchies (Ferro et al. 2010: 212; Ferro et al. 2011: 211). So, TSOMs are believed to provide an explanatory basis for both psycholinguistic and linguistic accounts of lexical parsability.
5.1 Temporal Self-Organising Maps

In the present dissertation, we use TSOMs to simulate dynamic effects of lexical storage, organization and competition. TSOMs, (Ferro et al. 2010; Ferro et al. 2011; Marzi et al. 2012; Pirrelli et al. 2014; Marzi et al. 2014; Marzi et al. 2015; Marzi & Pirrelli 2015; Marzi et al. 2016) are unsupervised artificial neural networks that learn to dynamically memorize input strings as chains of maximally-responding processing nodes (Best Matching Units or BMUs), whose level of sensitivity to input symbols in specific contexts is a continuous function of the distributional regularities of the input symbols during training. In a TSOM, each processing node has two layers of synaptic connectivity: an input layer connecting the node to the current input stimulus (e.g. the letter of a written word), and a (re-entrant) temporal layer connecting the node to all other nodes.

Given the BMU at time $t$, the temporal layer encodes the expectation of the current BMU for the node to be activated at time $t+1$. The strength of the connection between consecutively activated BMUs is trained through the following principles of correlative learning (compatible with Wagner & Rescorla (1974) equations). Given the input bigram $ab$, the connection strength between BMU of $a$ at time $t$ and BMU of $b$ at time $t+1$ will a) increase if $a$ often precedes $b$ in training (entrenchment), b) decrease if $b$ is often preceded by a symbol other than $a$ (competition).

The interaction between entrenchment and competition in a TSOM accounts for important dynamic effects of self-organization of stored words (Marzi et al. 2014, 2016). In particular, high-frequency words tend to recruit specialized (and stronger)
chains of BMUs, while low-frequency words are responded to by more blended (and weaker) BMU chains. In what follows, we report how well a TSOM can accommodate the complexity of the Greek verb system, by controlling factors such as word frequency distribution, degrees of inflectional regularity and word length.

In this dissertation, TSOMs are used to demonstrate the benefits as well as the requirements of the supported methodological approach, rather than show them as the best possible computational model in comparison to other models with similar properties (e.g. Li et al. 2007; Mayor & Plunkett 2010; Althaus & Mareschal 2013; among others). In fact, we delimit our study and focus on a specific neuro-computational model in order to present the points of our analysis more clearly and convincingly. As a result, most of the arguments and points claimed below will probably hold for other existing models with some qualifications (Marzi & Pirrelli 2015: 499).

5.2 Experimental evidence

The above-mentioned theoretical and psycho-linguistic analyses of MG conjugation offer evidence for degrees of morphological regularity, based on the interaction between formal transparency (graded levels of stem similarity) and (un)predictability of stem allomorphy. Hence, this evidence appears to question the sharp dichotomy that the dual mechanism claims for. In fact, no clear distinction between affix processing and allomorph retrieval can account for the interaction of formal transparency and predictability in MG word processing. In this chapter, we test the hypothesis that this evidence is compatible with a parallel processing architecture (a Temporal Self-organising Map) where processing and storage are in fact mutually implied.

It is important to note that the architecture of TSOMs aims to determine a general-purpose access level of lexical information, whose primary input is the most peripheral encoding of lexical forms. Thus, the emerging structures consist of purely formal redundancy in surface input data, without any information about the lexico-semantic content, morpho-syntactic or morpho-phonological features of words. Since in such a system, lexical organization is based on memory-based processing strategies, we can examine how several input factors affect acquisition and different processing tasks (Marzi & Pirrelli 2015: 525).

5.2.1 Dataset and training regimes

In setting my data, I used the corpus SUBTLEX-GR (BCBL 2016) constituted by Basque Center on Cognition, Brain and Language. The raw material comes from approximately 6,100 unique subtitle files, 4,001 corresponded to films and 1,507 to television series episodes (an average of 7.9 episodes per series). Accordingly, there are more than 27 million space-separated tokens of SUBTLEX-GR, 84.8% were taken from movies and only 15.2% from television shows. The films and the television series to which the subtitles corresponded were mostly USA productions (71.7%), in line with
the American dominance of the film and TV industry. Out of the remaining 28.3% subtitle files, nearly one third (9.5%) corresponded to UK productions and around two thirds (18.8%) to non-English productions, mostly French, German and Spanish. Hence, most of the subtitles used were transcripts of the English language.

In particular, here I used the restricted version of the corpus (SUBTLEX-GR_restricted.txt). This version resulted from cross-checking the corpus with a MG spell-checker which includes more than 1,600,000 inflected word forms (Symfonia software, ILSP). Through this process, spelling errors due to optical character recognition mistakes and word types not found in the Symfonia software were removed. The rejected strings made up a total of 75.6% of the types, but importantly, only 16.5% of the tokens. For the remaining 145,631 different word types, accounting for a total of 23,152,956 tokens, a set of additional frequency and lexical measures was also calculated. This version was preferred because it is much more manageable for linguistic research, since it includes only attested word tokens and types.

The training dataset contains present, simple past (aorist) and simple future indicative, active voice forms of 50 different verbs, for a total of 750 attested different verb forms (15 forms for each of the 50 most frequent paradigms), whose frequency distributions are sampled from the FREQcount section of the Greek SUBTLEX-GR corpus. From each paradigm, 15 inflected forms were extracted: the full set of present indicative (6) and simple past (6) forms and the singular forms of simple future (3). This ensures that training data will include all the cases of stem alternants (allomorphs and suppletive forms), since these paradigm cells mainly exhibit stem alternation in MG. Moreover, as I was interested in effects of global paradigm-based organization, forms were mostly selected from regular, formally transparent paradigms with active form. This means that I excluded paradigms with gaps in the tested tenses, impersonal and deponent verbs. Nonetheless, I included high-frequency paradigms with suppletive forms or/and non-predicted allomorphy (such as those of ‘vlepo ‘I see’, ‘leo ‘I say’ and katala’veno ‘I understand’) which were present in the training set (Ralli 2014). Concerning systematic free variants (e.g. mi’lo/mi’lao ‘I speak’), the most frequent form for each cell was selected, so as to avoid making a distinction between basic and alternative verb forms (Voga et al. 2012).

For this dissertation, I adopt the methodological, descriptive and explanatory adequacy resulting from the work of Marzi et al. (2014, 2016), Marzi and Pirrelli (2015) and Marzi et al. (2016). Following their previous experiments on German and Italian, the dataset was administered to a 42x42 node map for 100 learning epochs. In order to study the dynamics of word and paradigm acquisition, we trained identical maps on the same dataset administered under two distinct regimes. In one training regime (skewed distribution or SD), input forms were repeatedly presented to the map as a function (into the 1-1001 range) of their frequency distribution extracted by FREQcount section of the Greek SUBTLEX-GR corpus, for a total number of 29,464 token presentations per learning epoch. In the second training regime (uniform distribution or UD), the same set of input forms were shown to the map 5 times each, for a total number of 3,750 presentations per epoch. In both regimes, the order of word presentation was
randomized and results were averaged over five instances of the same map trained under each training regime for 100 epochs.

By varying frequency distributions and comparing the effects of inflectional complexity of training data on lexical access, word recall and word acquisition, we wanted to gain some knowledge about the interaction between morphological regularity and several other factors.

5.2.2 Word and paradigm acquisition

For the retrieval of the results, we tested the two groups of maps (uniform and skewed distributions) on the task of word recall, and compared their behaviour through the time course of lexical acquisition. Following Marzi et al.’s (2014) previous work, word recall is defined as the process of retrieving a word form from its chain of BMUs. Successful recall is possible if inter-node connections on the temporal layer are finely tuned to the distribution of symbols in the training data. The more accurate the re-entrant temporal coding is, the easier for the map to retrieve the symbols of a word in their appropriate order. Marzi et al. (2014) make the further reasonable assumption that a word is acquired by a TSOM when the map is in a position to recall the word accurately and consistently from its BMU chain. Errors occur when the map misrecalls one or more symbols in the input string, by either replacing them with different symbols or by outputting correct symbols in the wrong order. Partial recall is counted as an error as well. Average recall accuracy at epoch 100 turned out to be considerably high: 99.6% (std = 0.1%).

In order to understand the way that inflectional paradigms are acquired and the interplay between time of acquisition, frequency distribution, transparency and regularity, it is necessary to present some general trends about how these factors affect word acquisition. Hence, for the definition of the role of word frequency in word acquisition, we plotted the average frequency of correctly recalled words by learning epochs (Figure 2). The plot shows that skewed maps acquire the most frequent (probabilistically more supported) forms first, to acquire less supported sequences at later stages.

![Figure 2: Average frequency of correctly recalled words by learning epochs for skewed distributions](image)
The impact of this trend is even clearer when we compare the time course of lexical acquisition (recall accuracy) on uniform maps and skewed maps separately. In Figure 3, the plot shows how many words are acquired per learning epoch, as a percentage of all input words. The black solid line counts recall accuracy of word types in uniform maps, grey solid line types in skewed maps and dashed line tokens in skewed maps as well. Clearly, results are shown by both word types and word tokens for skewed maps only. Firstly, it is clear that TSOMs acquire word forms by token frequency, with higher-frequency words being quickly memorized and successfully recalled at earlier learning epochs. Moreover, skewed maps are found to acquire word types more slowly than uniform maps, suggesting that there is a significant advantage in having all forms presented an equal number of times during training. This advantage is eliminated in a map trained on skewed distributions since, as the training algorithm minimizes the probability for a map to make a recognition error, a SD map will tend to acquire the most frequent forms first, thereby neglecting less frequent ones (Marzi et al. 2014: 273-274; Marzi & Pirrelli 2015: 519). This is also shown by the mean frequency of correctly recalled forms at each epoch (Figure 2). Since top-frequency words are not necessarily the most similar to the majority of words in the training set (i.e. wordlikeness), using the memory resources for storing high-frequency items may effectively slow down lexical acquisition altogether (see also Ellis & Morrison 1998; Zevin & Seidenberg 2002).

![Figure 3: The time course of lexical acquisition (recall accuracy) on uniform maps (black solid line) vs. skewed maps for both type counts (grey solid line) and token counts (dashed line)](image)

This is expected, given the dynamics of selective specialization of TSOMs. A highly frequent input string tends to repeatedly activate the same pattern of nodes, strengthening the connections between sequentially activated BMUs, and establishing a dedicated, highly responsive IAP (Marzi et al. 2016: 91). Moreover, high token frequency is observed to avoid regularization effects through time (Bybee 1985, 1995; Corbett et al. 2001), and to leave deeply entrenched memory traces in the mental lexicon (Alegre & Gordon 1999; Baayen et al. 2007). This allows us to observe how paradigm regularity interact with word token frequency (see also Bybee 1985, 1995; Corbett et al. 2001).
To deal with this issue, for each paradigm we define its time of acquisition as the mean acquisition epoch of all forms belonging to the paradigm. The paradigm acquisition epoch provides an estimate of the average time it takes for all forms of the paradigm to be recoded in a time-sensitive way, for them to be recalled accurately from their corresponding activation chains (Marzi et al. 2014: 75). In Figure 4, we plotted the acquisition epoch of each Greek verb paradigm over both training conditions: skewed (black circles) and uniform (white circles). On the vertical axis, paradigms are arranged by increasing acquisition epochs in the skewed training regime, with cumulative paradigm token frequencies shown in brackets. This seems to suggest that some high-frequency paradigms are acquired very early because of the role of the (cumulative) token frequency of their members: however, unlike words, this is not always true.

![Figure 4: Acquisition time of paradigms ranked by increasing learning epoch for skewed (black circles) and uniform (white circles) distributions](image)

Firstly, the actual timing of paradigm acquisition appears to depend more on relative frequencies than on absolute frequencies. The vast majority of paradigms are acquired at a quicker rate in a uniform training regime than when they are presented with skewed frequencies: the number of epochs it takes to complete the acquisition of a paradigm after the first member of the paradigm is acquired (or paradigm acquisition span) is significantly shorter in uniformly-trained maps (UD < SD, p<.00005).

A notable exception to this general trend is represented by a few highly irregular paradigms (e.g. ˈleo ‘I say’, ˈvlepo ‘I see’, ˈpao ‘I go’) that are presented in training
with high cumulative token frequencies. Marzi et al. (2014: 277) account for this evidence suggesting that it highlights a non-trivial interaction between frequency and irregularity. More specifically, in highly-irregular paradigms with extensive unpredictable stem alternation, relatively isolated stems are acquired in a piecemeal, item-based fashion (see Bybee 1995). Unlike widely-distributed, more predictable stems, alternating stems are found in fewer cells of the paradigm and can take little (or no advantage) of cumulative frequency effects across cells. High token frequency can tip the balance in favor of early memorization, allowing irregular stems to successfully offset their low type frequency. This is what happens with ‘leo ‘I say’, ‘vlepo ‘I see’, ‘pao ‘I go’. In medium to low frequency irregular paradigms, however, this is more difficult to happen, and piecemeal memorization takes longer to set in. Moreover, the latter verbs in our dataset exhibit higher formal redundancy (and nesting) than the former (compare katala’veno ‘I understand’ ~ ka’talava ‘I understood’ vs. ‘perno ‘I take’ ~ ‘pira ‘I took’), allowing them to take advantage of uniform distribution as well. This memory effect may explain why the significant correlation between having all forms presented an equal number of times during training and learning epoch is limited to regular paradigms (UD < SD, p<.0005), where the aspectual marker and systematic or predictable stem allomorphy lead to more transparent shared stems, while it is weaker (UD = SD) with irregular paradigms where the stem transparency depends on the kind of non-systematic inherent allomorphy developed diachronically (Figure 5), suggesting once again that regularity in MG is not an all-or-nothing concept.

![Figure 5: Acquisition time of regular (left) and irregular (right) paradigms ranked by increasing learning epoch for skewed (black circles) and uniform distributions (white circles)](image)

Hence, token frequency effects are more hardly detectable when it comes to acquisition of regular paradigms, simply because predictable and systematic stems steadily appear in different forms of the paradigm and can take advantage of their cumulative frequency (Marzi et al. 2014: 277). According to Marzi et al. (2016: 94), regularity is observed to
favor acquisition of both high- and low-frequency words, due to the facilitatory effect of having more words that consistently activate the same pattern of shared nodes, creating partially overlapping or blended IAPs. Verb forms sharing the same stem tend to activate IAPs. Each time any of those forms is shown to the map, the connections between shared BMUs are strengthened over again. This prompts a boosting effect in acquisition, whereby a shared stem in a paradigm is responded to by a pattern of nodes whose level of entrenchment depends on stem family frequency rather than on word token frequency. Therefore, the skewed distribution does not facilitate the acquisition of regular paradigms (UD < SD, p<.0005). Especially, Figure 6 shows that this is true for both verb classes of regular paradigms; sigmatic with systematic stem alternants (UD < SD, p<.05) and sigmatic with phonological stem alternants (UD <SD, p<.001).

![Figure 6: Acquisition time of regulars with morphologically-conditioned stem alternants (left) and regulars with phonologically-conditioned stem alternants (right). Paradigms ranked by increasing learning epoch for skewed (black circles) and uniform (white circles) distributions](image)

In order to understand, however, why uniform distributions determine a clear advantage in time of acquisition, and why the advantage is unevenly apportioned between regular and irregular paradigms, as shown in Figures 5 & 6, and to shed light on this global interaction, we have to look at the way (ir)regularity interacts with other factors, affecting the developmental dynamics of memory organization. In terms of TSOMs, we had to examine the variable patterns of connectivity between recruited nodes, their levels of competition and co-activation on the map to understand the time course of paradigm acquisition and the degrees of perception of structure in the morphological lexicon. Following these lines, morphological structure can be investigated as the emergent property of a densely interconnected pool of nodes, whose global behaviour is governed by local patterns of connectivity (Marzi et al. 2014: 277).

The above evidence suggests that the interplay between frequency and word acquisition makes the actual timing of paradigm acquisition pretty complex. This is
because maps do not memorize words in isolation but in formally-related word families (or paradigms). The impact of word token frequencies on the time course of lexical acquisition, thus varies depending on the comparative amount of formal redundancy shared by input words (Marzi et al. 2014: 274).

Below, we account for the same evidence, by emphasizing the role of formal preservation of the stem (or stem transparency) in the paradigm as key to facilitation in morphological acquisition (and processing). Our data meet evidence from language acquisition and experimental psycholinguistics, underscoring the importance of formal redundancy for speakers’ perception of morphological structure (Tsapkini et al. 2002c: 116, 2004: 616; Stavrakaki & Clahsen 2009a: 117; Stathopoulou & Clahsen 2010: 872). Results were analysed using Linear Mixed Effects (LME) models with experiment repetitions and training items as random variables.

In Figure 7 we fitted the marginal plot of the interaction between word length and regular vs. irregular verb classes for MG, using a LME model fitting word learning epochs, with (log) word frequency, inflectional regularity and word length as fixed effects. This plot shows an interesting crossing pattern: shorter irregulars are acquired earlier than length-matched regulars of comparable frequency, but long irregulars are acquired later than long regulars.

\[ \text{GR skw: learnEp} = 1 + \logFreq \times \text{len} \times \text{IR} + (1)\text{repld} + (1)\text{lexExpId} \]

In order to clarify this interaction, we plotted a marginal plot of the interaction between (log) word frequency, length and regular vs. irregular verb classes, using a LME model fitting word learning epochs (Figure 8). Generally, it shows that high-frequency words are acquired more easily than low-frequency words. However, what is important here is that word length has different effects on the two classes of regularity only in the cases of the shortest and longest words, while the intermediate case of length (len=8.00) does not affect significantly differently the two categories. Furthermore, it is shown again
that irregular words (len=5.00) benefit from shorter stems and they are learnt earlier than the equivalent regulars, whereas this advantage is eliminated in the case of longer words (len=12.00), where regular forms are acquired earlier than irregulars.

Bompolas et al. (2016, submitted) account for this interaction, taking into consideration the extensive allomorphy exhibited by all Greek verb classes. Especially, in regular verbs, where perfective stem formation requires -s- suffixation that, subsequently, fires a phonologically-conditioned or morphologically systematic stem alternation, perfective stems are systematically longer than their imperfective counterparts, and are acquired after them. However, since imperfective stems are redundantly embedded in perfective stems, learning a long regular perfective form is easier (i.e. it takes a comparatively shorter time) than learning an irregular perfective form of comparable length. As a result, the kind of allomorphy characterizing a verb is crucial for its time acquisition; morphologically systematic and phonologically predictable stem allomorphy makes the stem of a word more transparent and facilitates the acquisition. Nevertheless, the beneficial effect of transparency has a more limited impact on learning, and is observed only for longer regular words. This is a regularity-by-transparency effect, and explains why long regular forms tend to be acquired (on average) more easily than long irregular forms.

### 5.2.3 Lexical access and word recall

To further investigate the impact of degrees of formal transparency on the processing of Greek verb forms, we account for this evidence in terms of a co-activation/competition dynamic in the tasks of serial word access, based on prediction, and word recall, based on parallel activation of target BMUs (Marzi et al. 2016).
Following Marzi et al.’s (2016: 89) previous work in serial word access, we simulate how the map can predict an incrementally presented input word. After training, each word in the training set is progressively presented to a TSOM by showing one symbol at a time on the input layer. Upon each symbol presentation, the TSOM is prompted to complete the current input string, by anticipating its possible continuation. In word recall, given a word’s IAP, we can use it as an input activation pattern to test how well the map can retrieve the word from its pattern. We simulate this by letting the map go through a word IAP, and iteratively output, at each time tick, the label of the current BMU.

![Figure 9: Marginal plot of interaction effects between log frequency (x axis), stem length (lenSt), suffix length (lenSx) and inflectional regularity in a LME model fitting Difficulty of Prediction (DoP) (y axis) by TSOM instances trained on Greek verb forms in the SD regime. Dotted line = irregulars (I), solid line = regulars with morphologically-conditioned stem alternation (RMA), dashed line = regulars with phonologically-conditioned stem alternation (RPA).](image)

At the first task we monitored the behaviour of Greek TSOMs on serial word access, by modelling how well an input word can be predicted by a map. So, we fitted a LME model of difficulty of word prediction (DoP) on Greek data (Figure 9), with (log) frequency and degree of regularity as fixed effects. The basic idea is that the more transparent the base form shared by the words of a paradigm, the harder the full word prediction on the grounds of competition for suffix selection. Generally, the plot shows that fully inflected irregular word forms are, on average, easier to be processed (and accessed) serially than regular words. Marzi et al. (2016: 96) account for this evidence on the basis of co-activation of stem sharing words which triggers competition for suffix selection; the more verb forms share the same stem, the stronger the competition for accessing an inflectional suffix. Suffixes in regular words are surrounded by many competitors and, thus, are less easy to predict than irregulars. In Greek regular verb forms, stem sharing and stem alternation leads to longer endings, which are distinguishable at later positions, increasing the amount of uncertainty at serial
processing in the selection of an upcoming symbol and the inflectional ending(s). On the other hand, in irregular paradigms, stem internal change and suppletion decrease the degrees of co-activation/competition for suffix selection. Significantly, the plot supports the assumption that paradigms exhibiting systematic stem allomorphy (e.g. *miˈl-o ‘I speak’ ~ ‘mili-s-a ‘I spoke’) are more transparent than those involving phonological stem allomorphs (e.g. *piˈstev-o ‘I believe’ ~ ‘pistep-s-a ‘I believed’). Regular word forms involving systematic stem-allomorphy contain more recurrent sub-lexical structures shared by other related forms (blended IAPs) than phonological stem allomorphy does, and, thus, cause effects of inhibition at predictability.

Turning to the second task, we modelled the difficulty of word recall (DoR) from IAPs. We conducted a LME analysis of the interaction between (log) word frequency, word length and inflectional regularity in word recall, whose marginal plot is shown in Figure 9. Unlike serial word access, which only relies on the map’s prediction bias for the most likely candidate symbol in a pool of competing candidates, recall is based on the integration of the map’s temporal expectation with the IAP of the target word to recall. As a general trend, when we control for (log) word frequency, regular verbs are found to be systematically easier at recall than the equivalent irregulars. This class of verbs are a highly entropic family whose members of the same stem family tend to develop blended activation patterns, and benefit from cumulative activation of more word types sharing the same stem, since there is no other expected family member but the target word to be recalled from its IAP (see also Marzi et al. 2016: 95).

Bompolas et al. (2016, submitted) interpret this effect once more as a regularity-by-transparency interaction. In particular, when we control for length (Figure 11), regulars with stem allomorphy make it comparatively easier to be recalled than RPA and irregulars, suggesting that this is an effect of stem transparency on cumulative
frequency. Consequently, the importance of accounting for degrees of transparency in MG stem formation is highlighted here: regular paradigms involving morphological allomorphs are the most systematic (and transparent), and are found to be easier to recall; phonological allomorphy is less systematic, and, thus, harder to recall; irregular paradigms, including non-systematic, formally more opaque stem alternation, are the hardest to recall. As shown by the difference in slope between the solid line and the other two lines of Figure 11, facilitation increases with word length, supporting our interpretation of the crossing pattern in the bottom panel of Figure 7 (see also Bompolas et al. 2016, submitted).

Figure 11: Marginal plot of interaction effects between length (x axis) and degrees of stem regularity in a LME model fitting DoR (y axis) by TSOM instances trained on Greek verb forms. Dotted line = irregulars (I), solid line = regulars with morphologically-conditioned stem alternation (RMA), dashed line = regulars with phonologically-conditioned stem alternation (RPA)

The finding (Figure 11) that stem-final systematic or predictable change, as in the case of regulars, leads to significantly easier recall than stem-internal changes and non-systematic/non-predictable stem-final changes, as is the case of irregulars, offers extra support to the assumption that the kind of stem allomorphy is crucial for the definition of the different levels of morphological regularity in MG. Significantly, this seems to involve a regularity-by-transparency interaction, with predictability playing a second role.

6. General discussion and conclusions

Among the different approaches to language acquisition and processing, the interactional ones are rendered as the most promising for the future (Chapter 1). In alignment with this inclination, the evidence provided in this dissertation focused on the definition of a performance-oriented notion of morphological regularity that might
overcome typical competence-based classifications presupposing a tight division between processing and storage. In this line, a paradigm-based approach of inflectional morphology was adopted, according to which the lexicon of a given language is a form-based (self-)organized network in which lexical structures are generated as a consequence of the paradigmatic relations among fully inflected word forms (Chapter 2). Analysis of MG conjugation provided by the seminal work of Ralli (1988, 2005, 2007) offers evidence which appears to question a dichotomous view of storage vs. rule-based processing mechanisms, because of the presence of a category of verb paradigms that combines both of these mechanisms (stored-allomorphy + rule-based suffixation) (Chapter 3). Indeed, the experimental analysis resulting from the Greek data supports the idea of degrees or gradations of morphological regularity, as a connectionist single-mechanism would predict, rather than qualitatively different operations based on the distinction between regulars and irregulars, as the dual-mechanism approach would predict (Chapter 4). Therefore, this case study focused on a distinguishing characteristic of MG conjugation: all verb paradigms, both regular and irregular ones, involve stem allomorphy in the formation of perfective verb forms. Hence, the difference between regular and irregular verbs could not be attributed to the categorical presence or absence of stem allomorphy as in the case with other languages, such as English, but to the type of stem allomorphy itself. In the remainder of this thesis, the hypothesis that this evidence is compatible with a parallel processing architecture (a Temporal Self-organising Map) where processing and storage are in fact mutually implied was tested (Chapter 5).

As for word and paradigm acquisition by Greek TSOMs (5.2.2), we observed a strong correlation between the amount of general vs. specialized resources that are apportioned by a TSOM through learning with sensitivity to a gradient of morphological regularity (see also Marzi & Pirrelli 2015; Marzi et al. 2014, 2016). More particularly, Marzi et al. (2016: 103-104) support that, regarding the more regular forms, they share redundant structures with other words to a considerable extent and, therefore, are stored in blended activation patterns and processed accordingly. Their acquisition is ultimately a function of how often these shared structures are found in input. Equivalently, in the case of Greek, the effect of word token frequency on entrenchment can capitalize on the cumulative token frequency of all members of the same word family, whose contribution is additive (Figures 4, 5 and 6). On the other hand, highly irregular forms are, by definition, relatively isolated, and get little (if any) support from the overall organization of the majority of lexical forms. Hence, they tend to develop IAPs that are rarely used elsewhere by a TSOM. However, irregular paradigms in MG are found to take advantage of both uniform and skewed distributions (Figure 5), suggesting that the members of some irregular paradigms share redundant structures to a considerable extent (regularity-by-transparency interaction).

Indeed, when effects of log frequency, word length and (ir)regularity are investigated (Figures 7 and 8), an interesting crossing pattern is observed: shorter irregulirs are acquired earlier than length-matched regulars of comparable frequency, but long irregulars are acquired later than long regulars. Bompolas et al. (2016, submitted) accounted for this evidence on the basis of the extended presentation of stem
allomorphy throughout Greek verb paradigms, no matter whether allomorphy is systematic, phonologically motivated or unsystematic. More specifically, in regular verbs, where perfective stem formation requires -s- affixation, perfective stems are systematically longer than their imperfective counterparts, and are acquired after them. Nonetheless, since imperfective stems are fully or partially nested in perfective stems, learning a long regular perfective form is easier (i.e. it takes a comparatively shorter time) than learning an irregular perfective form of comparable length. This is, again, a regularity-by-transparency effect, and explains why long regular forms tend to be acquired (on average) more easily than long irregular forms.

The further investigation of the impact of degrees of formal transparency on the processing of Greek verb forms was conducted by a LME analysis on the basis of word length, log frequency and regularity effects in the tasks of word recall and prediction (Chapter 5.2.3). In the case of word prediction (Figure 9), the general trend shows slower rates in processing regular verbs compared to irregular ones. According to Bompolas et al. (submitted), this happens on the grounds that stem allomorphy constrains the number of possible continuations across morpheme boundary, biasing the map’s expectations for suffix selection.

In the case of word recall, when we control for word length, forms in more regular paradigms are, on average, quicker to the recall than forms in irregular paradigms (Figure 10 and 11), since the former tend to cluster in larger word families, and this makes regulars more familiar or word-like, and their blended IAPs more routinized. On the contrary, forms in more irregular paradigms, suffer less from interference caused by co-activation of overlapping IAPs and, due to their higher frequency, they develop more deeply entrenched temporal connections (Marzi et al. 2016: 104). Interestingly, *Figures 10 and 11* demonstrate that facilitation of word recall increases with word length, supporting our interpretation of the crossing pattern of *Figure 7* (see also Bompolas et al. 2016, submitted).

A quantitative analysis of our experimental results accentuates a hierarchy of regularity-by-transparency impact on morphological processing. The evidence provided here underscores the role of formal preservation of the stem (or stem transparency) in the paradigm as a significant facilitator for morphological processing and acquisition. Our findings that fully transparent stems enable the initial processing of the word and rise perception of its morphological structure than more opaque stems do, are in agreement with a surface-oriented notion of morphological regularity, based on patterns of intra-paradigmatic formal redundancy. Furthermore, our results are consistent with psycholinguistic evidence in the same direction, laying emphasis on the significance of formal redundancy for speakers’ perception of morphological structure. Additionally, they are unfailingly compatible with psycholinguistic evidence of human processing, and seem to be consonant with research in Natural Morphology underlining the effects of iconic preservation of stem forms in regular inflection (Dressler 1996). This bears corroborating evidence to the conclusion that the type of stem allomorphy is what regulates the different levels of perceived morphological structure in MG, critically involving a regularity-by-transparency interaction, with predictability playing second fiddle.
The present analysis lays the foundations for a performance-oriented notion of inflectional regularity that may eventually cut across traditional dichotomous classifications. It is worth noting that dual-route models of lexical processing, which presume a sharp subdivision of work between storage and processing, are vitally based on a categorical, competence-based notion of morphological regularity fitting the inflectional systems of only some languages. Highly-inflecting languages such as MG appear to display a range of processes of stem formation that are noticeably more intricate and graded than traditional classifications are ready to admit. In turn, this level of complexity requires integrative, non-modular architectures of the human lexical processor. We put forward that TSOMs offer a promising implementation of such integrative architectures.

References


Robins, R. H. (1959) In defence of WP. *Transactions of the Philological Society 58*: 116-144.


