

For a Performance-oriented Notion of Regularity in Inflection: The Case of Modern Greek Conjugation

Stavros Bompolas*
Università di Patrasso

Marcello Ferro**
ILC-CNR

Claudia Marzi**
ILC-CNR

Franco Alberto Cardillo**
ILC-CNR

Vito Pirrelli**
ILC-CNR

Paradigm-based approaches to word processing/learning assume 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 equations that are key to modelling this complex dynamic at a considerable level of detail. We use these equations to simulate acquisition of Modern Greek conjugation, and we compare the results with evidence from German and Italian. 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.

1. Introduction

Issues of morphological (ir)regularity have traditionally been investigated through the prism of morphological competence, with particular emphasis on aspects of the internal structure of complex words (Bloomfield 1933; Bloch 1947; Chomsky and Halle 1967; Lieber 1980; Selkirk 1984, among others). Within this framework, one of the most influential theoretical positions is that morphologically, phonologically, and/or semantically transparent words are always processed on-line through their constituent elements, whereas irregular, idiosyncratic (non-transparent) forms are stored and retrieved as wholes in the lexicon (Pinker and Prince 1994). Likewise, Ullman and colleagues (1997) assume that the past tense formation of regular verbs in English requires on-line application of an affixation rule (e.g. *walk* > *walk+ed*), while irregular past tense forms, involving stem allomorphy (e.g. *drink* > *drank*), are retrieved from the lexicon.

Modern Greek introduces an interesting variation in this picture. First, stem allomorphy and suffixation are not necessarily mutually exclusive processes, but coexist

* Laboratory of Modern Greek Dialects, University of Patras, University Campus, 265 04 Rio Patras, Greece. E-mail: stavros.bompolas@gmail.com

** Istituto di Linguistica Computazionale "A. Zampolli", v. Moruzzi 1, Pisa, Italy.
E-mail: name.surname@ilc.cnr.it

in the same inflected forms (e.g. [ˈlin-o] ‘I untie’ > [ˈe-li-s-a] ‘I untied’, [aʝaˈp(a)-o] ‘I love’ > [aˈʝapi-s-a] ‘I loved’). Secondly, affixation rules may select unpredictable stem allomorphs: [aʝaˈp(a)-o] ‘I love’ > [aˈʝapi-s-a] ‘I loved’, [foˈr(a)-o] ‘I wear’ > [ˈfore-s-a] ‘I wore’, [xaˈl(a)-o] ‘I demolish’ > [ˈxala-s-a] ‘I demolished’. These cases suggest that inflectional (ir)regularity is not an all-or-nothing notion in Greek. Different inflectional processes may compound in the same words to provide a challenging word processing scenario (Tsapkini, Jarema, and Kehayia 2002b, 2004).

We offer here a computational simulation of the process of acquiring the Modern Greek verb system from scratch, based on exposure to fully-inflected forms only, with no extra morpho-syntactic or morpho-semantic information being provided. The idea is to investigate how aspects of morphological regularity can impact on early stages of word processing, prior to full lexical access. Our goal is to provide a causal model of the micro-dynamic of morphology-driven, peripheral processing effects, as observed in the experimental and acquisitional evidence of Modern Greek (see section 2). The simulation is implemented with a particular family of artificial neural networks, named Temporal Self-Organising Maps (TSOMs). Unlike traditional multi-layered perceptrons, TSOMs simulate the dynamic spatial and temporal organisation 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. Thus, TSOMs are ideal tools for assessing the differential impact of several aspects of regularity (from formal transparency, to predictability and word typicality) on the behaviour of a connectionist framework.

To anticipate some of our results, the paper provides a performance-oriented account of inflectional regularity in morphology, whereby perception of morphological structure is not the by-product of the design of the human word processor (purportedly segregating rules from exceptions), but rather an emergent property of the dynamic self-organisation of stored lexical representations, contingent on the processing history of inflected word forms, inherently graded and probabilistic. The evidence is in line with what we know of word processing by human speakers and illustrates the potential of a single, distributed architecture for word processing (Alegre and Gordon 1999; Baayen 2007), to challenge more traditional, modular hypotheses of grammar-lexicon interaction.

2. The evidence

Modern Greek conjugation, like Italian and unlike English, is stem-based. Each fully inflected verb form requires obligatory suffixation of person, number and tense markers that attach to either a bare or a complex stem in both regular ([aʝaˈp-o] ‘I love’ ~ [aˈʝapi-s-a] ‘I loved’) and irregular verbs ([ˈpern-o] ‘I take’ ~ [ˈpir-a] ‘I took’). Unlike English speakers, Greek speakers must always resort to an inflectional process to understand or produce a fully inflected form, no matter how regular the form is (Terzi, Papapetropoulos, and Kouvelas 2005, page 310).

Classifying a Greek verb as either regular or irregular thus requires observation of the stem formation processes whereby agreement and tense suffixes are selected. Accordingly, it is assumed that the presence or absence of the aspectual marker is a criterion for assessing the degree of regularity of a Greek verb. In particular, so-called “sigmatic” (from the Greek letter σ ‘sigma’) aorist (i.e. perfective) forms (e.g. [aʝaˈp-o] ~ [aˈʝapi-s-a]) are traditionally considered to be regular, in that they involve a segmentable marker (-s-) combined with phonologically predictable or morphologically

systematic stem-allomorphs. “Asigmatic” past-tense forms (e.g. [ˈpɛrn-o] ~ [ˈpɪr-a]), in contrast, exhibit typical properties of irregular inflection, since they involve unsystematic stem allomorphs (in some cases suppletive stems), and no segmentable affixes marking perfective aspect. This distinction has also been supported by psycholinguistic and acquisitional evidence (Stamouli 2000; Tsapkini, Jarema, and Kehayia 2001, 2002a, 2002b, 2002c, 2004; Mastropavlou 2006; Varlokosta et al. 2008; Stavrakaki and Clahsen 2009a, 2009b; Stavrakaki, Koutsandreas, and Clahsen 2012; Stathopoulou and Clahsen 2010; Konstantinopoulou et al. 2013, among others), leading some scholars to suggest that sigmatic past-tense forms are typically produced on-line by rules, and asigmatic forms are stored in, and accessed from the mental lexicon (Stavrakaki and Clahsen 2009a, 2009b; Stavrakaki, Koutsandreas, and Clahsen 2012).

However, careful analysis of the Greek verb system appears to question such a sharp processing-storage divide. In particular, Greek data provide the case of a mixed inflectional system where both stored allomorphy and rule-based affixation are simultaneously present in the formation of past tense forms. Ralli (1988, 2005, 2006) proposes a classification of verb paradigms based on two criteria; firstly, the presence vs. absence of the sigmatic affix and, secondly, the presence vs. absence of (systematic) stem allomorphy. As a result, we can define the following three classes of aorist formation processes (Tsapkini, Jarema, and Kehayia 2001, 2002a, 2002b, 2002c, 2004):

- (i) an affix-based class, requiring the presence of the aspectual marker *-s-*, and including verbs with a predictable phonological stem-allomorph (e.g., [ˈlɪn-o] ‘I untie’ ~ [ˈe-li-s-a], [ˈɣɾaf-o] ‘I write’ ~ [ˈe-ɣɾap-s-a] ‘I wrote’);
- (ii) a mixed class where active perfective past tense forms are produced by affixation of the aspectual marker *-s-* to a systematic morphological stem-allomorph (e.g., [miˈl-o] ‘I speak’ ~ [ˈmili-s-a] ‘I spoke’);
- (iii) an idiosyncratic verb class whose forms are based on either non-systematic stem-allomorphy (including radical suppletion), or no stem-allomorphy and no (sigmatic) aspectual marker (e.g., [ˈpɛrn-o] ‘I take’ ~ [ˈpɪr-a] ‘I took’, [ˈtro-o] ‘I eat’ ~ [ˈe-faɣ-a] ‘I ate’, [ˈkrin-o] ‘I judge’ ~ [ˈe-krin-a] ‘I judged’).

It should be noted that, in Greek regular verbs, transparency/systematicity and predictability are not mutually implied (Ralli 2005, 2006). The morphologically-conditioned allomorphy of class-(ii) verbs requires a systematic pattern of perfective stem formation, namely $X_{(a)} \sim X + V$ (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 Modern Greek free variant of the imperfective stem (e.g. [aɣaˈpo] ~ [aɣaˈpao]). The variable V in the perfective stem can be instantiated as *i*, *e* or *a*, and cannot be predicted, given the bare stem. On the other hand, the phonologically-conditioned allomorphs of class-(i) verbs (e.g. [ˈlɪn-] > [ˈe-li-s-]) are the outcome of exceptionless phonological rules, which nonetheless obfuscate a full formal correspondence (transparency) between the imperfective stem and the perfective stem.

Psycholinguistic experiments and evidence from language acquisition provide strong empirical support to the hypothesis that the human lexical processor is sensitive to the more nuanced classification of the Greek verb system reported above. Morphological regularity is not an epiphenomenon of the design of the human language faculty and the purported dualism between rule-based and memory-based routes. Rather, it is better understood as the graded result of the varying interaction of phonological, morphotactic and morpho-semantic factors. In particular, lack of full formal nesting between imperfective and perfective stems (as in [ðuˈlev-o] ‘I work’ ~ [ˈðulep-s-a] ‘I

worked') appears to have an extra processing cost for speakers. Tsapkini and colleagues (2002b) compared priming evidence obtained from Greek past-tense forms matched for orthographic overlap (50%) with present forms, in two different conditions of Stimulus-Onset Asynchrony¹ (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. [ðu'lev-o] ~ [ðulep-s-a], or [pɛrn-o] ~ [pir-a]) elicit significantly less priming than do fully-nested allomorphs (e.g. [mi'l-o] ~ [mili-s-a]), with intermediate cases of formal overlap (e.g. [lin-o] > [e-li-s-a]) 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 short and 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 shortened priming than orthographically opaque primes.

To sum up, analysis of Greek data offers evidence of graded levels of morphological regularity, based on the interaction between formal transparency (degrees of stem similarity) and (un)predictability of stem allomorphs. The evidence questions a dichotomous view of storage vs. rule-based processing mechanisms. In fact, no sharp distinction between affix processing and allomorph retrieval can possibly account for the interaction of formal transparency and predictability in Greek word processing.

A growing number of approaches to inflection, from both linguistic and psycholinguistic camps, developed the view that surface word relations represent a fundamental domain of morphological competence (Matthews 1991; Bybee 1995; Pirrelli 2000; Burzio 2004; Booij 2010; Baayen et al. 2011; Blevins 2016). Learning the morphology of a language amounts to acquiring relations between fully stored lexical forms, which are concurrently available in the speaker's mental lexicon and jointly facilitate processing of morphologically related forms through patterns of emergent self-organisation. This view presupposes an integrative language architecture, where processing and storage, far from being conceived of as insulated and poorly interacting modules, are, respectively, the short-term and the long-term dynamics of the same underlying process of adaptive specialisation of synaptic connections. Such an integrative architecture, upheld by recent evidence of the neuro-anatomical bases of short-term and long-term memory processes (Wilson 2001; D'Esposito 2007), crucially hinges on Hebbian principles of synaptic plasticity, which are, in turn, in keeping with mathematical models of discriminative learning (Rescorla and Wagner 1972; Ramscar and Yarlett 2007; Ramscar and Dye 2011; Baayen et al. 2011). The approach strikes us as particularly conducive to modelling the intricacies of Modern Greek inflection and provides an explanatory framework to account for human processing evidence. In what follows, we offer a connectionist implementation of this view.

1 Stimulus-Onset Asynchrony (SOA) is a measure of the amount of time between the start of the priming stimulus and the start of the target word. By varying SOA, one can elicit information about the time course of lexical access, namely the routes and procedures that are involved in earlier and later stages.

3. Computational Modelling

The advent of connectionism in the 80's popularised the idea that the lexical processor consists of a network of parallel processing units selectively firing in response to sensory stimuli. Arguably, the most important contribution of connectionism to the theoretical debate on lexical modelling at the time is that it explicitly rejected the idea that word recognition and production require a dichotomous choice between storage and processing. However, in spite of the *prima facie* psycho-computational allure of this view of the lexicon, early connectionist models also embraced a number of unsatisfactory assumptions about word learning and processing: from wired-in conjunctive coding of input symbols in context, to output supervision required by the gradient descent algorithm, to a model of word production as a derivational function mapping one lexical base onto fully inflected forms.

Later connectionist architectures have tried to address all these open issues. In particular, recurrent neural networks have offered a principled solution to (i) the problem of representing time, and (ii) the problem of learning without supervision. In simple recurrent networks (Jordan 1986; Elman 1990), the input to the network at time t is represented by the current level of activation of nodes in the input layer (as in classical connectionist networks) augmented with the level of activation of nodes in the hidden layer at the previous time tick ($t-1$). In this way, the network keeps track of its activation states and develops a serial memory of previous inputs.

Along the same lines, Temporal Self-Organising Maps (TSOMs) have recently been proposed to model the dynamic topological organisation of memory nodes selectively firing when specific symbols are input to the map in specific temporal contexts (Ferro, Marzi, and Pirrelli 2011; Marzi, Ferro, and Pirrelli 2014; Pirrelli, Ferro, and Marzi 2015). A temporal context is loosely defined as a temporal slot (position) in a time series of input symbols, or a window of surrounding symbols. Context-sensitive node specialisation is not wired in the map's connections at the outset (as in traditional connectionist models), but it is something that emerges as a function of input exposure in the course of training (Marzi et al. 2016). High-frequency input sequences develop deeply entrenched connections and highly specialised nodes, functionally corresponding to human expectations for possible continuations. Low-frequency input sequences tend to fire blended node chains, i.e. sequences of nodes that respond to a class of partially overlapping sequences. This is what distinguishes holistic, dedicated memorisation of full forms from chunk-based storage of low-frequency forms, sharing memory chunks with other overlapping forms (Marzi and Pirrelli 2015).

TSOMs offer an ecological way to conceptualise human word learning. As suggested by the psycholinguistic literature overviewed in section 2, children store all words they are exposed to, irrespective of degrees of regularity or morphological complexity. In addition to that, the self-organisation of items in the mental lexicon tends to reflect morphologically natural classes, be they inflectional paradigms, inflectional classes, derivational families or compound families, and this has a direct influence on morphological processing. In what follows, we provide a more formal outline of the architecture of TSOMs, to then explore their potential for modelling evidence from Greek inflection.

3.1 TSOMs

The core of a TSOM consists of an array of nodes with two weighted layers of synaptic connectivity (Figure 1). Input connections link each node to the current input stimulus

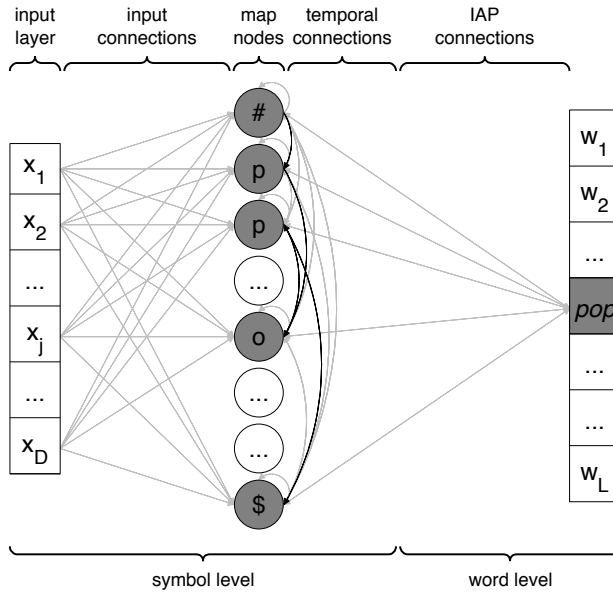


Figure 1
Overview of TSOM architecture.

(e.g. a letter or a sound), represented as a vector of values in the $[0, 1]$ interval, shown to the map at discrete time ticks. Temporal connections link each map node to the pattern of node activation of the same map at the immediately preceding time tick. In Figure 1, these connections are depicted as re-entrant directed arcs, leaving from and to map nodes. Nodes are labelled with the input characters that fire them most strongly. ‘#’ and ‘\$’ are special characters, marking the beginning and the end of an input word respectively.

Each time t a stimulus (e.g. an individual character or a phonological segment in a word) is presented in the input layer, activation propagates to all map nodes through input and temporal connections (short-term processing), and the most highly activated node, or Best Matching Unit (*BMU*), is calculated. Following this short-term step, node connections are made increasingly more sensitive to the current input symbol, by getting their weights $w_{i,j}$ (from the j -input value to the i -node) closer to the current input values $x_j(t)$. The resulting long-term increment $w_{i,j}$ is an inverse function $G_I(\cdot)$ of the topological distance $d_i(t)$ between the node i and the current $BMU(t)$, and a direct function of the map’s spatial learning rate $\gamma_I(E)$ at epoch E . $\gamma_I(E)$ is a dynamic parameter that decreases exponentially with learning epochs to define how readily the map can adjust itself to the input:

$$\Delta w_{i,j}(t) = \gamma_I(E) \cdot G_I(d_i(t)) \cdot [x_j(t) - w_{i,j}(t)] \tag{1}$$

Likewise, temporal connections are synchronised to the activation state of the map at time $t-1$, by increasing the weights $m_{i,h}$ (from the h -node to the i -node) on the connections between $BMU(t-1)$ and all other nodes of the map. The resulting long-term increment $m_{i,h}(t)$ is, again, an inverse function $G_T(\cdot)$ of their topological distance $d_i(t)$ from $BMU(t)$, and a direct function of the learning rate $\gamma_T(E)$ at epoch E :

$$\Delta m_{i,j}(t) = \gamma_T(E) \cdot G_T(d_i(t)) \cdot [1 - m_{i,h}(t)]; \quad h = BMU(t-1) \quad (2)$$

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 Rescorla-Wagner (Rescorla and Wagner 1972) 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:

1. increase if a often precedes b in training (entrenchment)
2. 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-organisation of stored words (Marzi, Ferro, and Pirrelli 2014; Marzi et al. 2016). In particular, high-frequency words tend to recruit specialised (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 learn the complexity of the Greek verb system, by controlling factors such as word frequency distribution, degrees of inflectional regularity and word length. Since our main focus here is on the dynamic of word processing, and on how this dynamic changes as the TSOM is exposed to more and more input words through learning, we will monitor both the developmental pace of acquisition (e.g. whether the map learns regulars more easily and quickly than irregulars or viceversa), and on the way the TSOM processes inflected forms at the final learning epoch. Other important issues, such as the ability to generalise to unknown forms (Marzi, Ferro, and Pirrelli 2014), or the ability to produce an inflected form on the basis of either a single uninflected base form (Ahlberg, Forsberg, and Hulden 2014; Nicolai, Cherry, and Kondrak 2015), or a pool of abstract morpho-lexical features (Malouf 2016) will not be addressed here.

4. The experiment

From the FREQcount section of the Greek SUBTLEX-GR corpus (Dimitropoulou et al. 2010), we selected the 50 top-ranked Greek paradigms by cumulative token frequency, and sampled 15 forms for each paradigm, for a total of 750 training items. For all 50 paradigms, forms were sampled from a fixed pool of paradigm cells: the full set of present indicative (6) and simple past tense (6) forms, and the singular forms of simple future (3). As we wanted to focus on effects of global paradigm-based organisation of active voice indicative conjugation, we excluded paradigms with systematic gaps, impersonal verbs, and deponent verbs. High-frequency paradigms with suppletive forms and/or non-systematic allomorphy (Ralli 2006, 2014) were included if they met our sampling criteria. Concerning systematic free variants (e.g. [aya'po] ~ [aya'pao]), the most frequent form for each cell was selected, to avoid making a distinction between basic and alternative verb forms (Voga, Giraud, and Anastassiadis-Symeonidis 2012).

The training dataset was administered to a 42x42 node map for 100 learning epochs. Upon each learning epoch, all 750 forms were randomly shown to the map as a function of their real word frequency distribution in the reference corpus, fitted in the 1-1001

range. To control for experimental variability, we repeated the experiment 5 times. After training, we assessed how well each map acquired the 750 input forms, using the task of word recall as a probe.

4.1 Word recall

In showing the string *#pop\$* one symbol at a time on the input layer (Figure 1), the activation pattern triggered by each symbol is incrementally overlaid with the patterns generated by all symbols in the string. The resulting integrated activation pattern (*IAP*) is shown in Figure 1 by levels of node activation represented as shaded nodes. Integrated activation is calculated with the following equation:

$$\hat{y}_i = \max_{t=1, \dots, k} \{y_i(t)\}; \quad i = 1, \dots, N \quad (3)$$

where i ranges over the number N of nodes in the map, and t ranges over the symbol positions in an input string k characters long. Intuitively, each node in the *IAP* is associated with the maximum activation level reached by the same node in processing the entire input word. Note that, in Figure 1, the same symbol p , occurring twice in *#pop\$*, activates two different *BMUs* depending on its position in the string. After presentation of *#pop\$*, integrated levels of node activation are stored in the weights of a third level of *IAP* connectivity, linking the map nodes to the lexical map proper (rightmost vector structure in Figure 1).

The resulting *IAP* is not only the short-term processing response of a map to *#pop\$*. The long-term knowledge sitting in its lexical connections makes the current *IAP* a routinized memory trace of the map processing response. In fact, a TSOM can reinstate the string *#pop\$* from its *IAP*. We call this reverse process of outputting a string from its *IAP* word recall.

The process consists in the following steps:

1. initialise:
 - (a) activate the word *IAP* on the map
 - (b) prompt the map with the start-of-word symbol #
 - (c) integrate the *IAP* with the temporal expectations of #
2. calculate the next *BMU* and output its associated label
3. if the end-of-word symbol \$ was not output:
 - (a) integrate the *IAP* with the temporal expectations of the *BMU*
 - (b) go back to step 2
4. stop

A word is recalled correctly from its *IAP* if all its symbols are output correctly in the appropriate left-to-right order.

It should be appreciated that, even when it is applied to the same word items used during training, word recall from the *IAP* is not a trivial process. Whereas in training each input stimulus is presented with explicit timing information (symbols are administered to the map one at a time), a word *IAP* is a synchronous activation pattern, where timing information is encoded only implicitly. In fact, accurate recall requires that the TSOM has developed a fine-grained association of map nodes with time events in

training, apportioning specialised time-bound nodes to systematically occurring input sequences. We can thus 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 *IAP*.

5. Data analysis

Average recall accuracy at epoch 100 turns out to be considerably high: 99.6 % (std = 0.1%). Results are analysed using Linear Mixed Effects (LME) models with experiment repetitions and training items used as random variables.

First, we compared the pace in the acquisition of Greek verb forms with the pace of acquisition of two other conjugation systems of comparable complexity but different language family: Italian and German. Results for Italian and German were obtained with the same training protocol used for Greek verbs (Marzi et al. 2016): 50 top-frequency verb paradigms were selected for each of the two languages, and the same pool of 15 forms was sampled from each selected paradigm. Input forms were administered for 100 epochs according to a function of their frequency distribution fitted in the 1-1001 range. Each training experiment was repeated 5 times, and results are averaged over all repetitions.

Figure 2 shows the marginal plot of the interaction between word length and regular vs. irregular verb classes for German, Italian and Greek, using an LME model fitting word learning epochs, with (log) word frequency, inflectional class and word length as fixed effects. In German and Italian, the distinction between regular and irregular paradigms is based on the criterion of absence vs. presence of stem allomorphy across all forms of a paradigm (Marzi et al. 2016). In Greek, we consider regular all paradigms showing a sigmatic perfective stem, and irregular those with an asigmatic perfective stem.

Unlike German and Italian (Figure 2, top and middle panels), where irregulars tend to be acquired systematically later than length-matched regulars are, and no significant interaction is found, Greek data (Figure 2, bottom panel) show 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.

Marzi and colleagues (2016) account for earlier learning epochs of both German and Italian regulars as an effect of stem transparency on cumulative input frequencies. With German and Italian regular verbs, stems are shown to the map consistently more often, since they are transparently nested in all forms of their own paradigm. This makes their acquisition quicker, due to specialised chains of stem-sensitive BMUs getting more quickly entrenched. Once a stem is acquired, it can easily be combined with a common pool of inflectional endings for tense and agreement, simulating an effect of (nearly) instantaneous (or paradigm-based, as opposed to item-based) acquisition. In contrast, Greek verb classes always present stem allomorphy throughout their paradigms, no matter whether allomorphy is systematic, phonologically motivated or unsystematic. In regular verbs, where perfective stem formation requires *-s-* affixation (verb classes *i* and *ii* above), 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 (verb class *iii* above). This is, again, a regularity-by-transparency

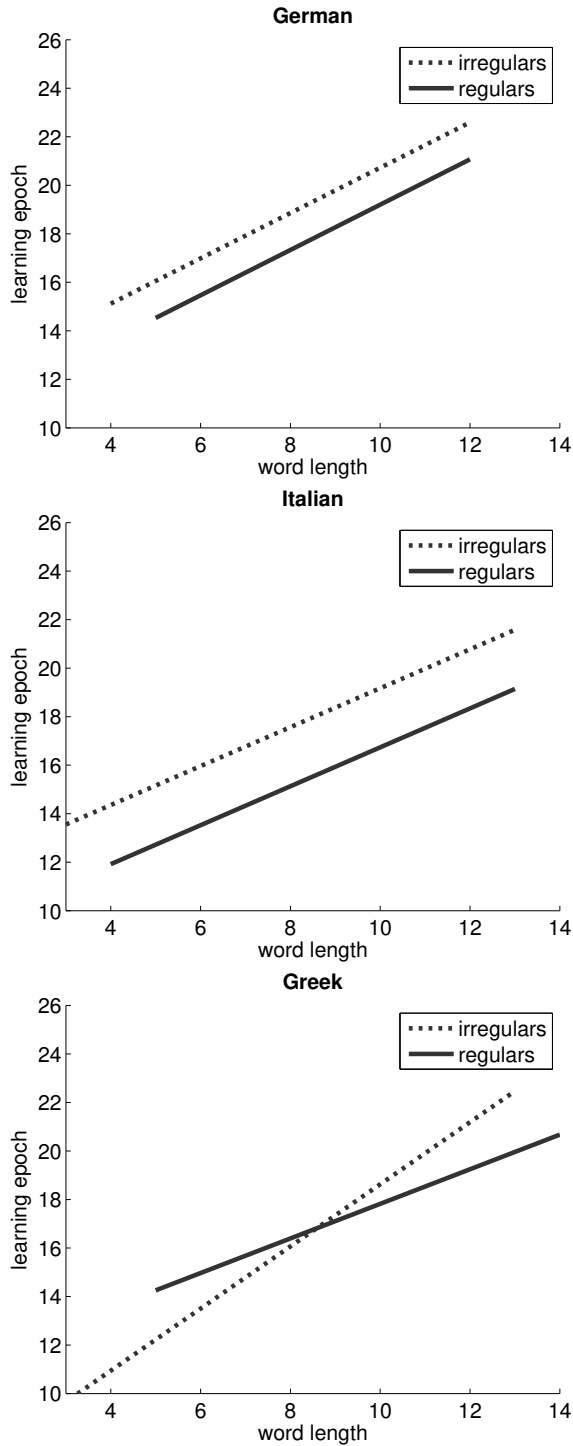


Figure 2 Marginal plots of interaction effects between word length, log frequency and inflectional regularity in an LME model fitting word learning epochs in German (top), Italian (middle) and Greek (bottom). Solid lines = regulars, dotted lines = irregulars. See main text for details on criteria for inflectional regularity in the three languages.

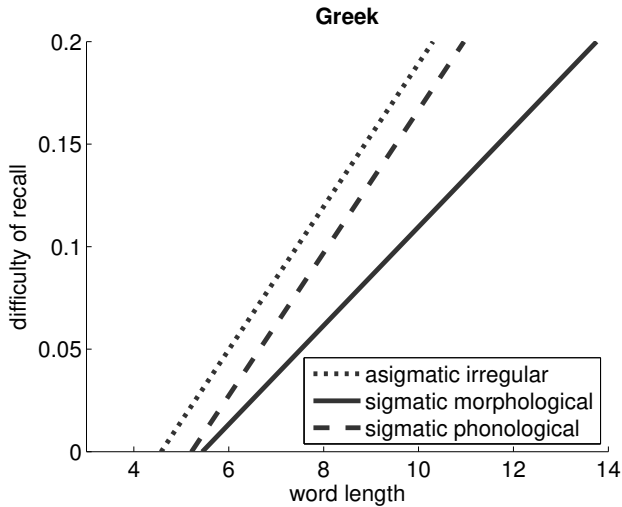


Figure 3

Marginal plot of interaction effects between word length (x axis), log frequency and degrees of stem regularity in an LME model fitting “difficulty of recall” (y axis) by TSOMs trained on Greek verb forms. Difficulty of recall is a direct function of the amount of noise filtering required for an input string to be correctly recalled from its *IAP*. Intuitively, the higher the level of activation of non target nodes in the *IAP*, the more difficult the recall of target nodes.

effect, and explains why long regular forms tend to be acquired (on average) more easily than long irregular forms.

To further investigate the impact of degrees of formal transparency on the processing of Greek verb forms, we conducted an LME analysis of the interaction between word length and classes of (ir)regularity in word recall (Figure 3). When we control for word length, regular verbs with sigmatic morphological allomorphs (e.g. [aʎa'p(a)-o] ~ [a'ʎapi-s-a], solid line in the plot) are recalled more easily than regular verbs with sigmatic phonological allomorphs (e.g. [ðu'lev-o] ~ [ðulep-s-a], dashed line in the plot). Difficulty of recall is estimated here as a direct function of the amount of filtering on node activation levels required for a word form to be recalled accurately from its *IAP* (Marzi et al. 2016). Intuitively, low levels of filtering mean that all relevant *BMUs* in the *IAP* are activated considerably more strongly than other irrelevant co-activated nodes, and are thus easier to recall. High levels of filtering suggest that the overall memory trace is difficult to recall due to the high activation of many spurious competitors. Note further that both regular classes are easier to recall than asigmatic irregular verbs (dotted line in the plot), which show, in most cases, formally more opaque allomorphs (e.g. [pern-o] ~ [pir-a]). As shown by the difference in slope between the solid line and the other two lines of Figure 3, recall difficulty increases with word length, supporting our interpretation of the crossing pattern in the bottom panel of Figure 2.

Finally, we assessed the role of predictability of both stems and affixes in Greek conjugation by classes of inflectional regularity. Figure 4 plots how easy it is for a map to predict an upcoming symbol at any position in the input string, given the preceding context. Our dependent variable, “ease of prediction” in the plot of Figure 4, is a function of how often a letter l , at distance k from the stem-ending boundary (to the left of the boundary for negative values, and to its right for positive values), is correctly predicted by the map. Note that ease of prediction is 1 if all letters at

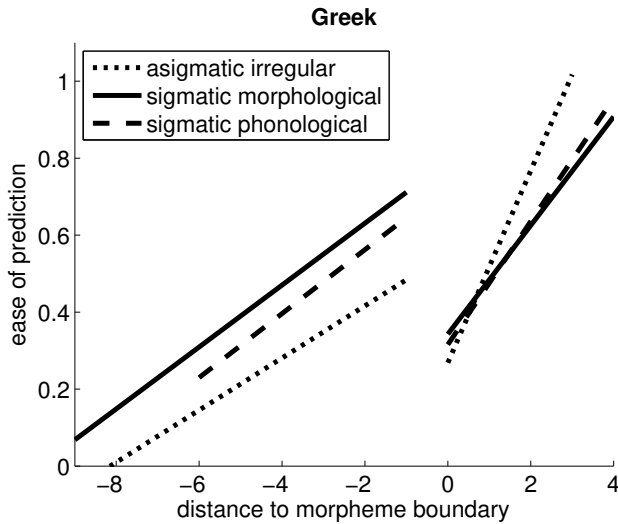


Figure 4

Marginal plot of interaction effects between the distance to the morpheme boundary (x axis, where $x = 0$ indicates the first symbol of the affix), and degrees of stem regularity in an LME model fitting “ease of prediction” (y axis) by TSOMs trained on Greek verb forms. Ease of prediction is 1 if all letters at position k from the morpheme boundary are correctly predicted by the TSOM, and it is 0 if no letters are predicted. The distance k has negative values if the letter precedes the boundary, and 0 or positive values if the letter follows the boundary.

position k are always correctly predicted, and equals 0 if no letter in that position is predicted. Intuitively, the slope of the marginal plot approximates the average number of prediction hits at any position in the input string.

In Figure 4, stem predictability and affix predictability are plotted as separate segments for the three classes of inflectional regularity, with $x = 0$ marking the first symbol after the base stem (e.g. the vowel i in [a'ʝapisa]). The general trend shows quicker rates in processing fully transparent stems compared with formally less transparent stems. Conversely, the slope of prediction scores for inflectional markers following transparent stems is less steep than the slope of inflectional markers following less transparent stems. It should be noted that the drop in prediction rate between the end of the stem and the start of the suffix in regular verbs gets smaller as we move to less transparent stem allomorphs. We take this level of uncertainty to be diagnostic for structural complexity: a drop in the level of expectation for an upcoming node at the morpheme boundary is the effect of a perceived morphological discontinuity on the temporal connectivity of the map. In keeping with distributional accounts of morphological constituency (Hay 2001; Bybee 2002; Hay and Baayen 2005, among others), this effect is also modulated by word frequency.

As expected, uncertainty is at its peak with fully transparent forms, due to the compounded effect of three factors: (i) nested base stems are shared by all inflected forms of a paradigm and are more easily predictable (due to entrenchment); (ii) for the same reason, they are trailed after by many different affixes; (iii) they undergo unpredictable perfective stem formation, which cannot easily be generalised across paradigms. Conversely, non transparent stems undergo several processes of stem alternation (either phonologically or morphologically motivated), and this makes it more difficult to predict their form during processing. However, uncertainty at the stem level

makes suffix selection easier across the morpheme boundary, when the map knows which stem allomorph occurs in input. In the end, stem allomorphy constrains the number of possible continuations across morpheme boundary, biasing the map's expectations. This processing dynamic accounts for an advantage in recalling verb forms with less transparent stems when these forms are comparatively shorter than those with more transparent stems (see Figure 2, bottom panel). The advantage progressively shrinks with word length, to become a disadvantage on longer forms.

6. General Discussion

Quantitative analysis of our experimental results highlights a hierarchy of regularity-by-transparency effects on morphological processing. In particular, the evidence offered here emphasises the role of formal preservation of the stem (or stem transparency) in the paradigm as a key facilitation factor for morphological processing.

Our case study focused on a distinguishing characteristic of Greek conjugation: all verb paradigms, both regular and irregular ones, involve stem allomorphy in past-tense formation. Hence, the difference between regular and irregular verbs could not be attributed to the categorical presence or absence of stem allomorphy as is the case with other languages, such as English and Italian (and, to a lesser extent, German). Rather, it should be attributed to the type of stem allomorphy itself. Our findings that fully transparent stems facilitate initial processing of the word and increase perception of its morphological structure than more opaque stems do, are in keeping with a surface-oriented notion of morphological regularity, based on patterns of intra-paradigmatic formal redundancy. In addition, they consistently meet psycholinguistic evidence of human processing, and appear to be in good accord with research in Natural Morphology laying emphasis on effects of iconic preservation of stem forms in regular inflection (Dressler 1996). This lends support to the conclusion that the type of stem allomorphy is what determines the different levels of perceived morphological structure in Modern Greek, crucially involving a regularity-by-transparency interaction, with predictability playing second fiddle.

The present analysis paves the way to a performance-oriented notion of inflectional regularity that may ultimately cut across traditional dichotomous classifications. It is noteworthy that dual-route models of lexical processing, which presuppose a sharp subdivision of work between storage and processing, crucially rely on a categorical, competence-based notion of morphological regularity fitting the inflectional systems of some languages only. Highly-inflecting languages such as Modern Greek appear to exhibit a range of processes of stem formation that are considerably more complex and graded than traditional classifications are ready to admit. In turn, this level of complexity calls for integrative, non modular architectures of the human lexical processor. We believe that TSOMs provide a promising implementation of such integrative architectures.

References

- Ahlberg, Malin, Markus Forsberg, and Mans Hulden. 2014. Semi-supervised learning of morphological paradigms and lexicons. In *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*, pages 569–578, Gothenburg, Sweden, 26-30 April, 2014.
- Alegre, Maria and Peter Gordon. 1999. Frequency effects and the representational status of regular inflections. *Journal of memory and language*, 40(1):41–61.
- Baayen, R. Harald. 2007. Storage and computation in the mental lexicon. In Gonia Jarema and Gary Libben, editors, *The Mental Lexicon: core perspectives*. Elsevier, Amsterdam, pages 81–104.

- Baayen, R. Harald, Petar Milin, Dusica Filipović Đurđević, Peter Hendrix, and Marco Marelli. 2011. An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological review*, 118(3):438–481.
- Blevins, James P. 2016. *Word and Paradigm Morphology*. Oxford University Press, Oxford.
- Bloch, Bernard. 1947. English verb inflection. *Language*, 23(4):399–418.
- Bloomfield, Leonard. 1933. *Language*. Henry Holt and Co., London.
- Booij, Geert. 2010. Construction morphology. *Language and Linguistics Compass*, 4(7):543–555.
- Burzio, Luigi. 2004. *Paradigmatic and syntagmatic relations in Italian verbal inflection*, volume 258, pages 17–44. John Benjamins, Amsterdam-Philadelphia.
- Bybee, Joan. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes*, 10(5):425–455.
- Bybee, Joan. 2002. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, 14(3):261–290.
- Chomsky, Noam and Morris Halle. 1967. *The Sound Pattern of English*. Harper and Row, New York.
- D'Esposito, Mark. 2007. From cognitive to neural models of working memory. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1481):761–772.
- Dimitropoulou, Maria, Jon Andoni Duñabeitia, Alberto Avilés, Jose Corral, and Manuel Carreiras. 2010. Subtitle-based word frequencies as the best estimate of reading behavior: the case of greek. *Frontiers in Psychology*, 1(218):1–12.
- Dressler, Wolfgang U. 1996. A functionalist semiotic model of morphonology. In Rajendra Singh, editor, *Trubetzkoy's Orphan: Proceedings of the Montréal Roundtable on "Morphonology: contemporary responses" (Montréal, October 1994)*, volume 144 of *Current Issues in Linguistic Theory*. John Benjamins Publishing Company, Amsterdam-Philadelphia, pages 67–83.
- Elman, Jeffrey L. 1990. Finding structure in time. *Cognitive Science*, 14(2):179–211.
- Ferro, Marcello, Claudia Marzi, and Vito Pirrelli. 2011. A self-organizing model of word storage and processing: implications for morphology learning. *Lingue e Linguaggio*, 10(2):209–226.
- Hay, Jennifer. 2001. Lexical frequency in morphology: is everything relative? *Linguistics*, 39(6):1041–1070.
- Hay, Jennifer B. and R. Harald Baayen. 2005. Shifting paradigms: gradient structure in morphology. *Trends in Cognitive Sciences*, 9(7):342–348.
- Jordan, Michael. 1986. Serial order: A parallel distributed processing approach. Technical Report 8604, University of California.
- Konstantinopoulou, Polyxeni, Stavroula Stavrakaki, Christina Manouilidou, and Demetrios Zafeiriou. 2013. Past tense in children with focal brain lesions. *Procedia-Social and Behavioral Sciences*, 94:196–197.
- Lieber, Rochelle. 1980. *On the organization of the lexicon*. Ph.D. thesis, MIT, Cambridge.
- Malouf, Robert. 2016. Generating morphological paradigms with a recurrent neural network. *San Diego Linguistics Papers*, (6):122–129.
- Marzi, Claudia, Marcello Ferro, Franco Alberto Cardillo, and Vito Pirrelli. 2016. Effects of frequency and regularity in an integrative model of word storage and processing. *Italian Journal of Linguistics*, 28(1):79–114.
- Marzi, Claudia, Marcello Ferro, and Vito Pirrelli. 2014. Morphological structure through lexical parsability. *Lingue e Linguaggio*, 13(2):263–290.
- Marzi, Claudia and Vito Pirrelli. 2015. A neuro-computational approach to understanding the mental lexicon. *Journal of Cognitive Science*, 16(4):493–535.
- Mastropavlou, Maria. 2006. *The effect of phonological saliency, and LF-interpretability in the grammar of Greek normally developing, and language impaired children*. Ph.D. thesis, Aristotle University, Thessaloniki.
- Matthews, Peter H. 1991. *Morphology*. Cambridge University Press, Cambridge.
- Nicolai, Garrett, Colin Cherry, and Grzegorz Kondrak. 2015. Inflection generation as discriminative string transduction. In *Proceedings of the Annual Conference of the North American Chapter of the ACL*, Denver, Colorado, USA, May 31 - June 5, 2015.
- Orfanidou, Eleni, Matthew H. Davis, and William D. Marslen-Wilson. 2011. Orthographic and semantic opacity in masked and delayed priming: Evidence from greek. *Language and Cognitive processes*, 26(4-6):530–557.
- Pinker, Steven and Alan Prince. 1994. Regular, and irregular morphology, and the psychological status of rules of grammar. In Susan D Lima, Roberta Corrigan, and Gregory K Iverson, editors, *The Reality of Linguistic Rules*. John Benjamins, Amsterdam, pages 321–351.

- Pirrelli, Vito. 2000. *Paradigmi in morfologia. Un approccio interdisciplinare alla flessione verbale dell'italiano*. Istituti Editoriali e Poligrafici Internazionali, Pisa.
- Pirrelli, Vito, Marcello Ferro, and Claudia Marzi. 2015. Computational complexity of abstractive morphology. In Matthew Baerman, Dustan Brown, and Greville G Corbett, editors, *Understanding and Measuring Morphological Complexity*. Oxford University Press, Oxford, pages 141–166.
- Ralli, Angela. 1988. *Eléments de la Morphologie du Grec Moderne: La Structure du Verbe*. Ph.D. thesis, University of Montreal.
- Ralli, Angela. 2005. *Morfologia [Morphology]*. Patakis, Athens.
- Ralli, Angela. 2006. On the role of allomorphy in inflectional morphology: evidence from dialectal variation. In Giandomenico Sica, editor, *Open Problems in Linguistics and Lexicography*. Polimetrica, Monza, pages 123–152.
- Ralli, Angela. 2014. Suppletion. In Georgios K Giannakis, editor, *Encyclopedia of Ancient Greek language and linguistics*. Brill, Leiden, pages 341–344.
- Ramsar, Michael and Melody Dye. 2011. Learning language from the input: Why innate constraints can't explain noun compounding. *Cognitive Psychology*, 62(1):1–40.
- Ramsar, Michael and Daniel Yarlett. 2007. Linguistic self-correction in the absence of feedback: A new approach to the logical problem of language acquisition. *Cognitive Science*, 31(6):927–960.
- Rescorla, Robert A. and Allan R. Wagner. 1972. A theory of pavlovian conditioning: variations in the effectiveness of reinforcement and non-reinforcement. In Abraham H. Black and William F. Prokasy, editors, *Classical conditioning II: Current research and theory*. Appleton-Century-Crofts, New York, pages 64–99.
- Selkirk, Elisabeth O. 1984. *Phonology and Syntax*. The MIT Press, Cambridge.
- Stamouli, Spyridoula. 2000. Simfonia, xronos ke opsi stin eliniki idiki glosiki diataraxi [agreement, tense, and aspect in specific language impairment in greek]. In *Proceedings of the 8th symposium of the Panhellenic Association of Logopedists*, Athens, Greece. Ellinika Grammata.
- Stathopoulou, Nikolitsa and Harald Clahsen. 2010. The perfective past tense in greek adolescents with down syndrome. *Clinical Linguistics & Phonetics*, 24(11):870–882.
- Stavrakaki, Stavroula and Harald Clahsen. 2009a. Inflection in williams syndrome: The perfective past tense in greek. *The Mental Lexicon*, 4(2):215–238.
- Stavrakaki, Stavroula and Harald Clahsen. 2009b. The perfective past tense in greek child language. *Journal of Child Language*, 36(01):113–142.
- Stavrakaki, Stavroula, Konstantinos Koutsandreas, and Harald Clahsen. 2012. The perfective past tense in greek children with specific language impairment. *Morphology*, 22(1):143–171.
- Terzi, Arhonto, Spyridon Papapetropoulos, and Elias D. Kouvelas. 2005. Past tense formation and comprehension of passive sentences in parkinson's disease: Evidence from greek. *Brain and Language*, 94(3):297–303.
- Tsapkini, Kyrana, Gonia Jarema, and Eva Kehayia. 2001. Manifestations of morphological impairments in greek aphasia: A case study. *Journal of Neurolinguistics*, 14(2):281–296.
- Tsapkini, Kyrana, Gonia Jarema, and Eva Kehayia. 2002a. A morphological processing deficit in verbs but not in nouns: A case study in a highly inflected language. *Journal of Neurolinguistics*, 15(3):265–288.
- Tsapkini, Kyrana, Gonia Jarema, and Eva Kehayia. 2002b. Regularity revisited: Evidence from lexical access of verbs and nouns in greek. *Brain and Language*, 81(1):103–119.
- Tsapkini, Kyrana, Gonia Jarema, and Eva Kehayia. 2002c. The role of verbal morphology in aphasia during lexical access. In Elisabetta Fava, editor, *Clinical Linguistics: Theory and Applications in Speech Pathology and Therapy*, volume 227. Benjamins, Amsterdam-Philadelphia, pages 315–335.
- Tsapkini, Kyrana, Gonia Jarema, and Eva Kehayia. 2004. Regularity re-revisited: Modality matters. *Brain and language*, 89(3):611–616.
- Ullman, Michael T., Suzanne Corkin, Marie Coppola, Gregory Hickok, John H. Growdon, Walter J. Koroshetz, and Steven Pinker. 1997. A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience*, 9(2):266–276.
- Varlokosta, Spyridoula, Anastasia Arhonti, Loretta Thomaidis, and Victoria Joffe. 2008. Past tense formation in williams syndrome: evidence from greek. In Anna Gavarró and Maria João Freitas, editors, *Proceedings of GALA 2007*, pages 483–491, Cambridge.

- Voga, Madeleine, Hélène Giraudo, and Anna Anastassiadis-Symeonidis. 2012. Differential processing effects within second group modern greek verbs. *Lingue e Linguaggio*, XI(2):215–234.
- Voga, Madeleine and Jonathan Grainger. 2004. Masked morphological priming with varying levels of form overlap: Evidence from greek verbs. *Current Psychology Letters: Behaviour, Brain & Cognition*, 2(13):1–9.
- Wilson, Margaret. 2001. The case for sensorimotor coding in working memory. *Psychonomic Bulletin & Review*, 8(1):44–57.