This is a contribution from *Specific Language Impairment. Current trends in research*.
Edited by Stavroula Stavrakaki.

The relationship between SLI in English and Modern Greek

Insights from computational models of language acquisition

Themis Karaminis [a, b] and Michael S. C. Thomas [a]
[a] Developmental Neurocognition Lab, Birkbeck /
[b] Centre for Research in Autism and Education, Institute of Education

We present a computational modelling approach to the study of SLI in two languages with different typological characteristics, namely English and Modern Greek. Our modelling approach was based on the development of three neural network (connectionist) architectures, each assumed to underlie the acquisition of a core domain of language (inflectional morphology, syntax comprehension, and syntax production). The architectures were exposed to artificial linguistic environments reflecting the characteristics of their target domains in English and Greek. Computational simulations also considered conditions of atypical learning constraints, corresponding to different theoretical proposals for the type of deficit underlying SLI. The simulation results, combined with some shared properties of the three models, point to a unified explanation of the impairment under the connectionist framework.

Introduction

Recent approaches to language acquisition have argued that empirical phenomena in language development could be accounted for in terms of an interaction between a language learning mechanism and the properties of the linguistic environment to which it is exposed (Hutzler et al. 2004; Mareschal et al. 2007a, b; Thomas & Karmiloff-Smith 2003; Seidenberg 2011). It is a challenge to characterise the nature of these interactions in greater detail. Addressing this challenge involves explaining phenomena in the acquisition of different languages, and moreover doing so for both typical development (TD) and atypical development

DOI 10.1075/lald.58.07kar

(AD). If successful, an important advantage of this approach would be its generality, in two forms: across languages and across language groups.

In this chapter, we present a study attempting a general approach to language acquisition based on computational modelling. Computational modelling is a powerful tool in cognitive science (Norris 2005; Plaut 2000; Thomas & Karmiloff-Smith 2002). It can be used to better understand theories of cognitive development, by elucidating the mechanisms that underlie various empirical effects and if possible, unifying diverse empirical effects via a single mechanistic implementation. Computational models are also useful for testing the ability of different theories to account for given phenomena. Finally, computational models advance theories by delineating their implications and generating novel testable predictions.

The particular type of computational modelling used in our study involves simulations with the connectionist or (artificial) neural networks (Rumelhart, McClelland & the PDP Research Group 1986; see also Karaminis & Thomas 2012a, b). Neural networks are a class of artificial learning systems inspired by the biology of the brain (Rumelhart, Hinton & McClelland 1986). They are based on architectures of neuron-like processing units interconnected via weighted connections and performing computations in a massively parallel fashion (Parallel Distributed Processing, PDP; Rumelhart, Hinton & McClelland 1986). These architectures can exhibit learning, when a learning algorithm (e.g., back-propagation of error; Rumelhart, Hinton & Williams 1986) shapes the weights between the processing units appropriately, during repeated presentations of a training set. For example, the training set might constitute a collection of stimuli representing the environment in which the child is brought up or the cognitive domain that is to be learnt (target domain). Artificial neural networks are therefore ideally suited to model interactions between learning systems and cognitive environments. More generally, neural network modelling is a research tool that conforms to the major theoretical commitments of the neuroconstructivist framework (Mareschal et al. 2007a, b; Thomas & Karmiloff-Smith 2003), such as the idea that theories of cognitive development should be consistent with theories of functional brain development; that development is an emergent phenomenon involving the interaction of experience-dependent learning systems with structured physical and social environments; and that this process is multiply constrained.

Within connectionist psycholinguistics, the focus is on learning of the structures of language. A fundamental contrast between connectionist accounts of language development and certain linguistic theories (e.g., Chomsky 1965, 1981, 1986, 1991, 1995, 1998; Pinker 1999) concerns the necessity of innate grammatical knowledge to acquire language. In Linguistic Theory, such knowledge translates to a processing structure involving explicitly defined rules operating on symbols. In connectionist models of language development no such rules are specified.

Mental representations are distributed and these give rise to rule-following behaviour, with regularity being the product of similarity-based processing of different stimuli within the PDP learning system (e.g., Plunkett & Marchman 1991, 1993; Rumelhart & McClelland 1986). In artificial neural networks, linguistic competence is an emergent behaviour.

The emergence of linguistic knowledge in connectionist architectures is an important finding by itself; neural network models, however, are more informative when used to identify assumptions and constraints under which architectures acquire their target domains in a psycholinguistically-plausible manner. They can address questions such as the following: (a) Which architectural assumptions entail that the learning in these models is symptomatic of child development, e.g., in terms of developmental error patterns and rates thereof? (b) How are statistical properties of the linguistic environment, as encoded in the training set of the model, reflected in the developmental trajectory? (c) (How) does the model account for the acquisition of different languages? (d) What about atypical language development? What sorts of constraints generate atypical linguistic profiles, such as those found in specific language impairment (SLI; Leonard 1998); (e) Where different patterns are found across languages could the same atypical constraints generate different patterns in different languages? These questions represent the main research aims of our computational study, as well as progress in the pathway towards a general account of language development under the neuroconstructivist framework.

Our research design involved the development of three neural network architectures, each assumed to underlie the learning of a core domain of language, namely inflectional morphology (IM), syntactic comprehension (SC), and syntactic production (SP). We assumed that the acquisition of IM involves learning to modify word stems appropriately according to grammatical context; that the acquisition of SC requires learning to assign thematic roles (e.g., 'agent', the noun performing an action denoted by a verb) to words of sentences, and that the acquisition of SP is the developmental process of learning to map meanings to particular linguistic forms. The three neural network architectures were exposed to training sets based on artificial languages that reflected the characteristics of their target domains, either in English or in Modern Greek. This cross-linguistic paradigm considered the contrast between two language typologies: a language with a fairly simple morphological system, wide use of morphologically unmarked forms, and a statistically-dominant subject-verb-object (SVO) word order; and a language with a fusional and complex system of morphology, absence of morphologically unmarked forms and a more flexible word order. To address the acquisition profile of SLI we considered theoretically-driven processing constraints based on different aetiological hypotheses for the impairment (see below). The focus was on the extent to which competing hypotheses on the nature of the underlying deficit of SLI could account for the profile of the impairment in both languages.

This chapter presents an overview of the modelling endeavour, focusing on the findings which are more relevant to the study of SLI (research questions (d) and (e)). We begin with a sketch of the target domains in the two languages and a description of the main types of deficits characterising their acquisition by children with SLI. These deficits defined the target empirical phenomena for the models. We next outline different theoretical accounts of SLI, and how these accounts have been explained and tested in previous connectionist models of SLI. We then present the three new models focusing on three key elements of each: (1) the cognitive architecture proposed to underlie the acquisition of the target domain; (2) the cross-linguistic differences between the training sets in the English and the Modern Greek version; and (3) the conditions of constrained processing found to generate profiles symptomatic of SLI in the two languages. The reader who is not interested in the modelling details might want to skip this section. In our final discussion, we outline some shared properties of the three models, and reflect on how these properties, considered together with combined results from simulations, point to a unified account of SLI under the connectionist framework.

Cross-linguistic differences between English and Modern Greek

Inflectional Morphology

English employs a fairly simple morphological system (e.g., Fromkin, Blair & Collins 2002), presenting a limited number of inflections and abundance of morphologically unmarked or default forms (e.g., base forms of verbs). Inflected forms are realised by combining unmarked forms with suffixes (e.g., progressive: -/in/), which may also distinguish allomorphs, i.e., phonological variants of the same morpheme (e.g., past tense allomorphs: -/t/, -/d/, -/^d/). Inflections are fully regular (e.g., progressive) or quasi-regular (e.g., past-tense). Quasi-regularity (e.g., Marslen-Wilson & Tyler 1998; Thomas & Karmiloff-Smith 2003) refers to the presence of a dichotomy between inflected forms that conform to the suffixation rule (e.g., talk/talked) and certain clusters of the inflected forms realised through alternative rules (e.g., past tense; identity: set/set, hit/hit; vowel change: know/knew, flow/flew).

The system of IM in Modern Greek is much more complex (Holton, Mackridge & Philippaki-Warburton 2003; Stephany 1997). Content words are obligatorily marked for a range of grammatical features. For example, nouns, which have grammatical gender, are inflected with respect to case and number; verbs are inflected with respect to person and number, as well as tense, aspect, and voice.

This variety of grammatical features is realised fusing word stems with appropriate suffixes, and in some cases infixes and prefixes. Unmarked forms of content words do not exist. Another important difference is the absence of a clear-cut dichotomy between regular and irregular inflection. This is because the language presents alternatives (conjugational classes) for the realisation of the different grammatical features, and these are often combined together resulting in further complexity (Stephany 1997: 185).

Syntax

English relies upon word order to convey meaning. The Subject-Verb-Object (SVO) word order is the statistically dominant pattern (e.g., Dick et al. 2001). A possible explanation for the reliance of English upon word-order is the limited use of morphological marking in the language. For example, morphology could aid understanding the meaning of active sentence only when the subject and the object were in different numbers (Subject-Verb agreement; see Dick et al. 2001).

Modern Greek, on the other hand, marks words according to their syntactic role (subject: nominative; object: accusative; Stephany 1997). There are degradations in the extent to which morphological marking distinguishes between nominative and accusative forms of nouns, with nominative and accusative forms being identical for neuter nouns (salience of case-marking; see Karaminis 2012: 365; also Lupyan & Christiansen 2002). Morphological marking, nevertheless, may be used for understanding the meaning of sentences and this allows for a more flexible order of constituents in sentences (cf. Holton et al. 2003: 426). In sentences with neutral intonation, such as those presented at the beginning of discourse (Holton et al. 2003: 427), the SVO and the VSO word-orders are considered as statistically dominant (Kail & Diakogiorgi 1994). However, it is not the case that the variations of word-order are pragmatically equivalent (Tzanidaki 1995). As Kail and Diakogiorgi (1994: 9) discuss, 'word order is constrained by some syntactic factors such as sentence type (dependent vs. independent), length and complexity of constituents and mainly by pragmatic factors (topic and focus)'.

Target empirical phenomena for SLI in English and Modern Greek

Extensive listings of findings from behavioural studies of SLI are provided in Bishop (1997) and Leonard (1998). Here, we focus on the data that were used to characterise the target empirical phenomena for the three models. Data on IM refer to the types of inflections in which deficits are more pronounced in SLI and the relevant error types that are produced. Data on SC refer to a specific psycholinguistic task on the interpretation of different syntactic structures (Dick et al. 2004). Data on SP refer to a task measuring accuracy rates and error patterns in the production of Wh-Questions (Stavrakaki 2006; van der Lely & Battel 2003).

Inflectional morphology

Children with SLI acquiring English as a first language present lower rates of correct responses than TD children in elicited production tasks for example, where a child is required to alter a target sentence into the past tense ('Every day I run to the shops. Yesterday I ____ to the shops.'). However, the impairment is not uniform across different inflections. For example, it is more pronounced in the production of the simple past tense than in the production of present progressive (Rice 2000; van der Lely & Ullman 2001). Compared to TD children, children with SLI produce inflection omission errors (e.g., Yesterday, I *eat* a candy) in notably higher rates, a phenomenon which has also been described under the Extended Optional Infinitive account of SLI (Rice et al. 1995). According to this account, the Optional Infinitive stage, an early developmental stage in which children produce unmarked forms in contexts that require morphological marking, is protracted in SLI. An Extended Optional Infinitive stage could suggest underspecified tense representations in SLI (Rice et al. 1995). Alternative explanations, however, consider the role of high frequency of base forms in English or the phonological similarity between base and inflected forms (Matthews & Theakston 2006). Other characteristics of morphological development in SLI include increased frequency effects for both regular and irregulars, and lower rates of rule-based inflection of novel verbs (wug/wugged), as well as overgeneralisation errors (e.g., eated) in irregular inflection (e.g., van der Lely & Ullman 2001). Finally, some phonological effects appear to influence morphology. The rates of correct responses in regular forms are conditioned by whether the word endings produced by the addition of the inflectional suffixes to base forms conform (e.g., keep/kept) or do not conform (e.g., merge/merged) to the phonotactics of the language (Marshall & van der Lely 2006). The rates are significantly higher for past tense forms that are phonotactically legal in SLI; a similar effect was also observed in TD children, but it was not statistically reliable (Marshall & van der Lely 2006).

Children with SLI acquiring Modern Greek present difficulties in the use of certain grammatical categories such as the masculine gender, the plural number, and the genitive case (Mastropavlou 2003a,b). These grammatical categories are also hard to learn for TD children (Mastropavlou 2003a,b; Stephany 1997). For both TD children and children with SLI, the main error pattern is substitutions from more frequent categories (e.g., using suffixes from the neuter gender to mark masculine nouns, Mastropavlou 2003a,b; Stephany 1997).

Stavrakaki, Koutsandreas and Clahsen (2012) examined the acquisition profile of children with SLI in the perfective past tense, used to denote events that were completed (perfective aspect) in the past. Although this domain presents multiple conjugational categories, it is possible to consider a fundamental distinction between verbs which require – among other features – the addition of an aspectual marker (-s-) to the verb stem to form the perfective past tense (henceforth: the sigmatic class, s: sigma) and verbs which do not require an aspectual marker (henceforth: the non-sigmatic class). The addition of the aspectual marker can be seen as a rule involved in the perfective past tense formation. The presence of such a rule is also supported psycholinguistically. Both TD children and children with SLI produce sigmatic responses for novel rhymes in elicited production tasks and present higher rates of overapplication of sigmatic responses in existing non-sigmatic verbs than vice versa (Stavrakaki & Clahsen 2009; Stavrakaki et al. 2012). Children with SLI, however, rely to a lesser extent upon the sigmatic rule than TD children. This is evidenced by lower rates of overapplication of the sigmatic past tense in the non-sigmatic category, higher rates of non-sigmatic responses in the sigmatic categories, and lower rates of sigmatic responses for novel rhymes (Stavrakaki et al. 2012).

For other grammatical features, evidence of the linguistic profile of children with SLI is inconsistent. For example, as the marking of the perfective past tense in verbs is obligatorily realised in conjunction with the marking of person and number through suffixation, studies have examined impairments in Subject-Verb (S-V) agreement in SLI, and whether these are more pronounced than impairments in tense marking. The case studies of Tsimpli (2001) and Clahsen and Dalalakis (1999) suggested that deficits of children with SLI in S-V agreement marking are more severe than deficits in tense-marking. This was based on the observation of many person substitution errors, which occurred with the perfective past-tense tense being marked correctly. The majority of these errors were overapplications of 3rd person singular forms (the most frequent person-number combination) to other persons, especially the 2nd person, which rhymes with it. Other studies, however, have found less severe impairments in the marking of person (Diamanti 2000; Smith 2008; Stamouli 2000).

Syntactic comprehension

Our target data for the abilities of English children with SLI on SC came from an on-line sentence interpretation task in which children were asked to identify the agent ('Whodunnit?'; Dick et al. 2004) in four types of sentences: Active Voice

(AV, 'The dog chases the cat'), Passive Voice (PV, 'The cat is chased by the dog'), Subject Cleft (S-Cl, 'It is the dog that chases the cat'), and Object Cleft (O-Cl, 'It is the cat that the dog chases'). The pattern for accuracy rates of children with SLI in the four types of sentences was AV = S-Cl>PV>O-Cl, suggesting ceiling accuracy rates for sentences presenting the statistically dominant agent-verb-patient[1] word order and lower rates for sentences with alternative word orders. Compared to TD children matched on language age, accuracy rates of children with SLI were significantly lower only for O-Cl. Finally, a significant (positive) effect of S-V agreement on sentence disambiguation was found only for O-Cl sentences.

In a similar acting-out task for Modern Greek (Stavrakaki 2002), which allowed the interpretation of the same types of sentences to be studied, the pattern for accuracy rates of both children with SLI and TD children was AV = S-Cl>PV>O-Cl, and again this difference coincided with ceiling accuracy rates in sentences with the most frequent agent-verb-patient word order and lower rates in sentences with other word orders. A difference from the English data (Dick et al. 2004) was that accuracy rates of Greek children with SLI were significantly lower than rates of TD children not only for O-Cl, but also for PV sentences. It should be noted that the data of Stavrakaki (2002) on Modern Greek did not consider effects of S-V agreement (studied for English SLI in Dick et al. 2004) and salience of case marking in SC in SLI. As we shall see, the Modern Greek version of the SC model generated predictions regarding these issues.

Syntactic production

Data for SP in SLI came from a study by van der Lely and Battel (2003). This study examined the production of subject and object Wh-Questions (e.g., who-Subject: 'Who frightened Mrs Brown?'; which-Object: 'Which cat did Mrs White stroke?'). Children with SLI presented lower accuracy rates in all types of questions than TD children matched on language-age, and this difference was more pronounced in Object questions, which employ unusual word orders (SV or OSV word orders). Differences were also observed with regards to error patterns. Children with SLI produced responses such as 'Who did Mrs Scarlett saw somebody?' denoting difficulties in employing the auxiliary combined with difficulties in suppressing the

1. We refer to the structure of sentences using the semantic terms *agent* and *patient* rather than the syntactic terms subject and object. Note, however, that nouns carrying the semantic role of the agent and patient also carry the syntactic role of subject and object in AV, S-Cl, and O-Cl sentences. The reverse holds for PV sentences (subject = patient, object = agent). For this reason, we will also use the term subject-verb number agreement, to refer to the rule according to which the subject and the verb should be in the same number.

presence of an object at the end of the sentence (gap-filling error, see van der Lely & Battel 2003). As van der Lely and Battell (2003: 170) discuss, the presence or the absence of these errors classified the children participating in the study as normal or impaired with 90% accuracy.

In a similar study of Wh-question production in Modern Greek, Stavrakaki (2006) reported that Greek children with SLI presented lower accuracy rates in all types of Wh-questions than TD children matched on language age. Similarly to the English data, the number of errors was greater in object questions. In terms of error patters, an important difference between children with SLI and TD children was the production (by the former) of responses that were grammatically unacceptable in the language. As discussed earlier, in Modern Greek the subject of sentences is always marked in the nominative case and the object in the accusative. An example of ungrammatical responses of children with SLI in Stavrakaki (2006) was the production of sentences in which both nouns were marked in the accusative case.

Theoretical accounts of SLI

For a comprehensive discussion of theoretical accounts of SLI see Bishop (1997), Leonard (1998), and Ullman & Pierpont (2005). For the purposes of this chapter, we will restrict ourselves to sketching the main similarities and differences between distinct theoretical proposals. A relevant issue is a fundamental distinction between two schools of thought in addressing the strongly language-specific profile of the impairment. One school of thought attributes the language-specific profile of SLI to a deficit in the brain systems involved in the processing of language, which can be directly characterised in linguistic terms. Such accounts (e.g., Gopnik 1990; Gopnik & Crago 1991; Rice et al. 1995; van der Lely 1996) offer a static view of the language system and typically propose that certain grammatical features or operations are absent from the start state or have not developed in the language of children with SLI. As these language-specific accounts of SLI are stated in terms of rules and operations on symbolic structures, they are predicated upon symbolic accounts of language (e.g., Chomsky 1981).

An opposing school of thought, known as non-language-specific accounts of SLI, proposes that the linguistic profile of SLI stems from impairments which are not localised in the linguistic-system. These accounts suggest that general processing limitations or slower processing (e.g., Bishop 1994; Kail 1994; Leonard, Bortolini, Caselli, McGregor & Sabbadini 1992) could affect the learning of language in SLI. A challenge for these accounts is to explain the differential impact of these limitations on language skills. A slightly different view considers deficits that are peripheral to the language system, in the sense that they are not localised in grammar. Such deficits are presented in speech perception (Tallal & Piercy 1973a,b), or in phonological working memory (Gathercole & Baddeley 1990). Non-language-specific accounts of SLI do not rely on the assumption of symbols and rule-like operations. Such accounts emphasise the impact of general constraints on a processing system, rather than particular anomalies on symbolic operations. In particular, these accounts posit that different aspects of language might be affected differentially, depending on their processing requirements (e.g., the surface hypothesis; Leonard 1989; Leonard, McGregor & Allen 1992). In that sense, non-language-specific accounts of SLI are compatible with the idea of a PDP system employing distributed representations.

More recent theoretical views of SLI present features of both language-specific and non-language-specific explanations. For example, the Procedural Deficit Hypothesis (PDH) (Ullman & Pierpont 2005) suggests that SLI stems from a deficit in the Procedural Memory System. This is a general brain system, underlying many linguistic and non-linguistic functions. With regards to language, the Procedural System supports the learning of rule-based operations. Thus, the deficit in the Procedural System affects symbolic operations involved in language processing.

Previous computational models of SLI

There are relatively few connectionist models of SLI and, importantly, they have been limited to the study of the acquisition of English. In that sense, an advance of the current study is addressing the profile of the impairment in a non-English language. Hoeffner and McClelland (1993) and Joanisse (2004) simulated deficits in morphological acquisition implementing a phonological/ perceptual deficit in neural network architectures acquiring aspects of English Inflectional Morphology. Thomas (2005) discussed the main assumptions of the Procedural Deficit Hypothesis (PDH) (Ullman & Pierpont 2005) for the impairment, considering different manipulations of the parameter space of a three-layered network learning the English past tense. In the domain of syntax, Joanisse and Seidenberg (2003) demonstrated that a phonological deficit could address deficits of children with SLI in anaphoric resolution. Thomas and Redington (2004) addressed deficits of children with SLI in complex sentence interpretation; this model has been extended in the current study to include morphological characteristics thereby becoming a more general model that is also applicable to Modern Greek. To our knowledge, there are no connectionist studies addressing Syntax Production in SLI.

Modelling Inflectional Morphology

The IM model, like the subsequent two models, used a 2x2 design with the factors of language and language group. The three models needed to simulate developmental data for both English and Modern Greek, and for both typical and SLI groups. In addition the neural network model of IM was novel in its implementation of a generalised inflectional system, which considered multiple grammatical classes (nouns, verbs, and adjective) and multiple inflections within a grammatical class (e.g., noun: base forms, plural, genitive). To do so, this model brought together elements of several previous models (e.g., multiple inflections within a grammatical class: Hoeffner & McClelland 1993; lexical-semantics information: Joanisse & Seidenberg 1999; multiple grammatical classes: Plunkett & Juola 1999; three-layered feed-forward neural network trained on an artificial language: Plunkett & Marchman 1991, 1993; Thomas & Karmiloff-Smith 2003). This resulted in the architecture shown in Figure 1. Four types of information were presented in the input layer: the phonological form of a base form or a word stem for Modern Greek, its lexical semantics, the grammatical class it belonged to, and the type of morphological modification to be performed (target inflection). The network should use this information to produce the appropriate inflected form (its phonological form) and the output layer. An overarching assumption was that this system was embedded within other systems providing the relevant information to drive inflection; and that the goal of this system was to the phonological form of words that was conditioned to the grammatical context in which it was to appear (see Karaminis & Thomas 2010).

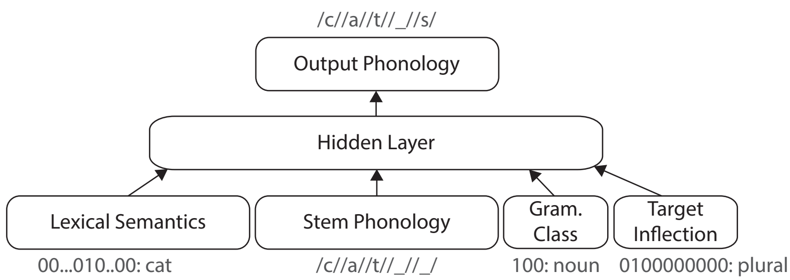

Figure 1. The three-layered feed-forward architecture of the model of IM (Gram.Class: Grammatical Class)

The architecture of the model was exposed to two training sets, one for the English and one for the Modern Greek version. The two training sets reflected the particularities of the domain of IM in each language. The morphological complexity of Modern Greek was reflected in a notably larger collection of input-output mappings (26,400 vs. 5,600 in the English version). More forms were required in Modern Greek in order to accommodate the greater variety of grammatical features and inflections, as well as the presence of conjugational classes. Both training sets were also constrained by assumptions on the frequencies of different inflections, regular and irregular classes, allomorphs, or grammatical classes. These were based on measurements of language corpora (English: tagged Brown Corpus; Francis & Kucera 1982; Modern Greek: Hellenic National Corpus; Hatzigeorgiou et al. 2000). In the case of Modern Greek, where a tagged corpus was not available, frequency assumptions were also based on descriptions in other studies (e.g., Stephany 1997).

The starting point to simulate SLI was predicated upon the ability of a *default* or 'TD' version of the model to learn the target domains in the two languages and to simulate a range of phenomena characterising morphological development in English and Modern Greek. This was verified through qualitative and quantitative comparisons with the corresponding empirical data (for details, see Karaminins & Thomas 2010). The model demonstrated that this was achieved through the flexible integration of the different cues. For example, lexical-semantics were particularly important for the learning of irregular inflections (see also, Joanisse & Seidenberg 1999); grammatical class information was not as important in English (it was redundant because the 'target inflection' cue could also be used to determine whether a word was a noun, a verb or an adjective), but was essential for the learning of Modern Greek IM.

In line with the research design, we then considered various conditions of atypical learning constraints, introduced in the parameter settings of the model at the onset of training. Examples of such constraints were the use of fewer units in the hidden layer of the network[2] (see also, Thomas & Redington 2004), a low-discriminality sigmoid activation function[3] (see also Thomas 2005), or phonological representations of reduced strength (see also Hoeffner & McClelland 2003). These constraints were motivated by theoretical positions on the aetiology of atypical language development. For example, phonological representations of reduced

2. There are usually no theoretical claims for the inclusion of layers of hidden units in neural network architectures. However, layers of hidden units are in many cases essential for architectures to exhibit learning of their target domains. As Plunkett and Marchman (1993: 28) discuss, hidden units provide neural networks with `the capacity to construct internal representations of the input vectors… [and] …introduce a non-linear component into the mapping process'. In that sense, the use of fewer hidden units in a neural network architecture corresponds to a reduction in the processing power of the network.

3. A low-discriminality sigmoid activation is typically implemented using a lower 'temperature' value, which reduces the steepness of the curve (for details, see Thomas 2005; Thomas & Karmiloff-Smith 2003). This manipulation effectively results in the processing units being less sensitive to variations in the strength of incoming activation.

strength could correspond to theoretical accounts suggesting that a phonological/perceptual deficit underlies atypical language acquisition.

In the English version of the model, the atypical condition that provided the best fit to the data involved the combination of two deficits: processing limitations through fewer units in the hidden layer and a perceptual deficit through phonological representations of reduced strength. The same type of deficit also allowed the model to capture empirical effects on SLI in Modern Greek.

In certain cases, it was possible to evaluate the ability of the model to fit the empirical data based on quantitative comparisons. Details of such comparisons are shown in Figures 2 and 3. Figure 2 refers to the English past tense, based on the dichotomy between regular and irregular verbs. Figure 3 focuses on the 2nd person singular of the perfective past tense in Modern Greek, based on the distinction of the so-called sigmatic ('regular') and non-sigmatic ('irregular') categories (Stavrakaki & Clahsen 2009; Stavrakaki et al. 2012). The formation of the perfective past tense in the sigmatic category involves, among other features, the addition of an aspectual marker -s- to the stem (s: sigma, in Modern Greek). This rule does not apply to the non-sigmatic category, where perfective past-tense formation involves idiosyncratic stem changes. The comparisons addressed the extent to which the TD and the SLI versions of the model differed in the same way as TD children and children with SLI in the empirical data. This was evaluated by calculating correlation coefficient values between two vectors, one comprising the rates of different types of responses (correct responses and main error types) in the TD and the SLI models and one comprising the rates of the same type of responses in the empirical data. The modelling data were obtained after matching the TD model and the data of TD children on correct responses in the regular past tense or the sigmatic perfective past tense. The correlation coefficient value was 0.52, $p = 0.08$ in the case of the English past tense and 0.78, $p < 0.001$ in the case if the perfective past tense in Modern Greek. The correlation coefficient values were taken to suggest a qualitative fit between the model and the data for the English model, and a quantitative fit for the Modern Greek model. The quantitative comparisons between the model and the data were complemented by qualitative comparisons regarding specific response types, based on visual inspection. Visual inspection of Figure 2 suggests that the English version of the model was successful in capturing the drop of accuracy rates in regular and irregular past tense and the increase in the rates of omission errors; however, it was unsuccessful in producing overgeneralisation errors in higher rates. The Modern Greek version (Figure 3) was successful in simulating the drop of performance and the production of 'other' responses (e.g., imperfective past tense forms) in increased rates in SLI; however, it underestimated the rates of correct responses and sigmatic responses in the non-sigmatic category. Finally, the Modern Greek version of the model also generated a testable prediction regarding the production of 3rd person singular forms in a perfective past tense inflection task examining S-V agreement, in TD and especially in SLI. Note that these responses not included in Figure 3, as the empirical data did not address S-V agreement. Further details on this prediction are provided in Discussion section.

Modelling Syntactic Comprehension

Unlike the model of IM, the model of SC was of more limited scope. It extended the model of Thomas and Redington (2004) to account for the interaction of morphological and word-order information in sentence interpretation, and to account for Syntactic Comprehension in typical development and SLI for Modern Greek. The architecture of the model is shown in Figure 4. It was based on a Simple Recurrent Network (SRN, Elman 1990, 1991, 1993), a type of neural network architecture with a feedback mechanism that allows some sort of memory of previous activation states to develop (cf. the role of context layers below). In the SC model, sentences were presented in the input layer in a word-by-word fashion. Each word was presented at a separate processing cycle (timestep) using information on its morphological structure (base form and morphological suffixes) and prosodic information, i.e., word length and stress. The network also employed a context layer keeping a copy of the activation of the hidden layer in the previous cycle multiplied by a recurrency factor *mu* (parameter of the model). The aim of this layer was to provide the network with information on its internal state, as conditioned by words that had been processed in the previous timesteps. It corresponded to a type of working memory. The network should output the next word in a sequence (a standard task for SC in SRN networks, see Elman 1991, 1993), whilst also setting to the correct value a unit encoding the relative order of the agent and the patient in the sentence.

The broader theoretical framework underlying the proposal of this architecture was the Competition Model (MacWhinney & Bates 1989). According to this theory, sentence interpretation is driven by the integration of different cues present in sentences, which could be local (e.g., morphological marking of words) or more global (e.g., word-order). Cues vary in strength, e.g., the presence of the word *by* in a sentence is a strong cue for a passive in English (see Dick et al. 2004; Thomas & Redington 2004). Different cues might perform in competition or in coalition in sentence disambiguation. An example of competition is between the presence of a noun-verb-noun word order in PV, which could point to the statistically dominant agent-first interpretation, and the presence of the word *by*, suggesting the patient-first word order of PV. An example of coalition is when the agent

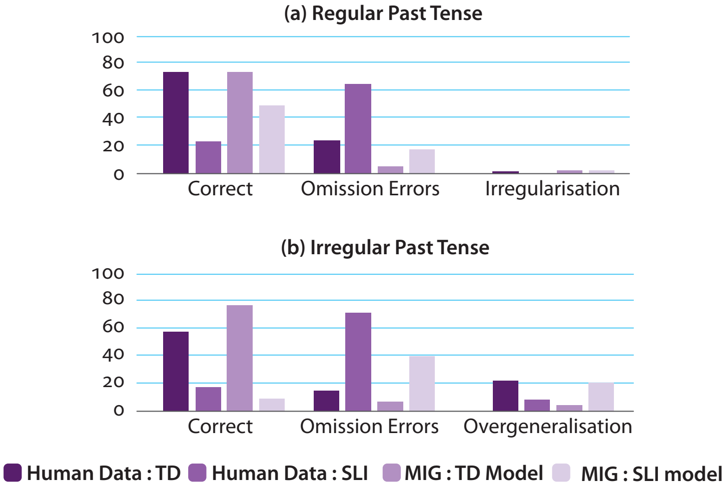

Figure 2. The acquisition profile of the English past tense in human data for TD and SLI and the model of IM. (The model was matched to the children data based on accuracy rates on regular verbs in TD)

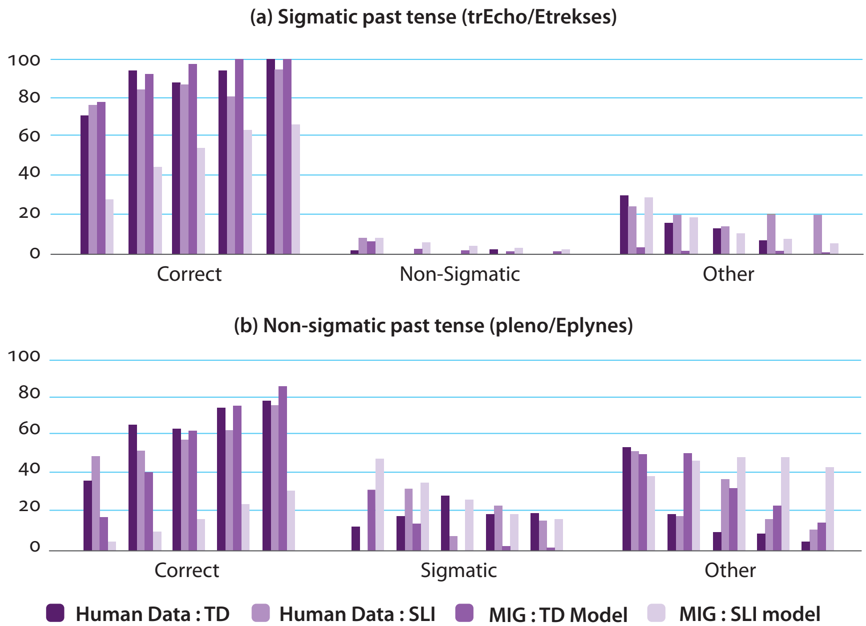

Figure 3. The acquisition profile of the 2nd person singular of the perfective past tense in empirical data for TD and SLI and the model of IM. (The model was matched to the children data based on accuracy rates on sigmatic verbs in TD)

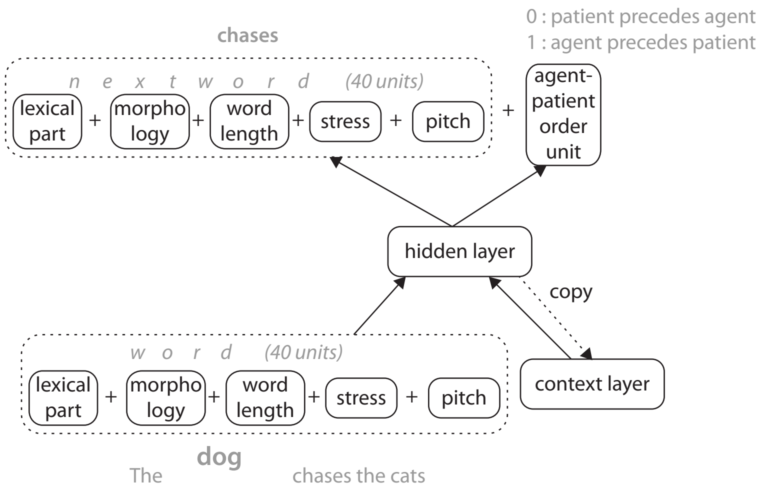

Figure 4. The architecture of the SC model. The figure refers to the English version and shows the presentation of the active sentence 'The dog chases the cats'

and the patient are in different numbers, where S-V agreement could be combined with other cues (e.g., word order) for sentence interpretation.

The model of SC was exposed to an artificial grammar of simple sentences and a frequency structure for the different sentence types (Thomas & Redington 2004). Constructions included declarative, interrogative and imperative sentences. These were general sentence types, which were divided further in subtypes (e.g., declarative: intransitive, active, passive, object cleft, subject cleft). The grammar was implemented twice, i.e., in the English and the Modern Greek version and following the particularities of each language (English: plural marking; Modern Greek: number marking and case marking of nouns, with degradations in salience of case-marking). Note, however, that the model relied crucially upon the assumption that the frequencies of different sentence types would follow child language in English and Modern Greek. However, it also simplified the range of alternative word-orders that were possible in Modern Greek.

The default version of the model could integrate frequencies of different types of sentences with the cues present in them and identify the order of the agent and patient in them. An inspection of the response of the network across sentence presentation at different stages of training revealed a transition from an early phase where the decision was driven by local cues to a later phase where more global cues were taken into account. This transition determined a pattern of accuracy rates in different types of sentences that was symptomatic of the patterns in the empirical data for English (Dick et al. 2004) and Modern Greek (Stavrakaki 2002).

The condition that simulated SLI in the two languages was the use of a lower recurrency factor. This condition effectively reduced the salience of cues related to long-distance dependencies in the network. It conforms to recent ideas that syntactic deficits in SLI may arise from deficits in phonological working memory (Richardson, Thomas & Price 2010; see also Joanisse & Seidenberg 2003). Interestingly, in the model, the reduced working memory condition captured the different pattern of dissociations between SLI and TD in the two languages (English: differences only in O-Cl; Modern Greek: differences in PV and O-Cl). The reason why accuracy in PV sentences was affected more severely in the Modern Greek version of the model was the use of case-marking (exclusively in this version). PV sentences employed a pattern for the marking of case (first noun: nominative; last noun: accusative) that suggested an incorrect agent-first interpretation. The strength of this cue was increased by the lowering of the recurrency factor. Another interesting result was the prediction that, given that the strength of cues related to word order was attenuated, children with SLI should rely more upon morphological cues for sentence interpretation. Further, as morphological marking prompted an incorrect interpretation of PV sentences and a correct interpretation of O-Cl, salient patterns of case marking would impede the comprehension of the former and facilitate the interpretation of the latter.

Modelling Syntactic Production

The architecture of the model of SP is shown in Figure 5. The model was based on a production-SRN architecture (Chang, Dell, Bock & Griffin 2000) using a recurrent neural network which was a somewhat reversed version of the model of SC. The architecture should produce a sentence in a word-by-word fashion at the output layer. It was assumed that the intention of the speaker to describe an event with a given type of sentence entailed the presence of the following two types of information (see Chang et al. 2002), which were presented in the input layer: a message (e.g., agent: dog; action: chase; patient: cat); and an event semantics, characterising surface similarities and differences between syntactic structures (e.g., which-Object question: has agent, has patient, interrogative, wh-question). The input layer also included information on the words that had been produced in recent timesteps, as well as a context layer providing a copy of the activation of hidden layer in the previous timesteps. These allowed the network to track where in the sentence its word-by-word output had reached, and corresponded to a form of working memory for production.

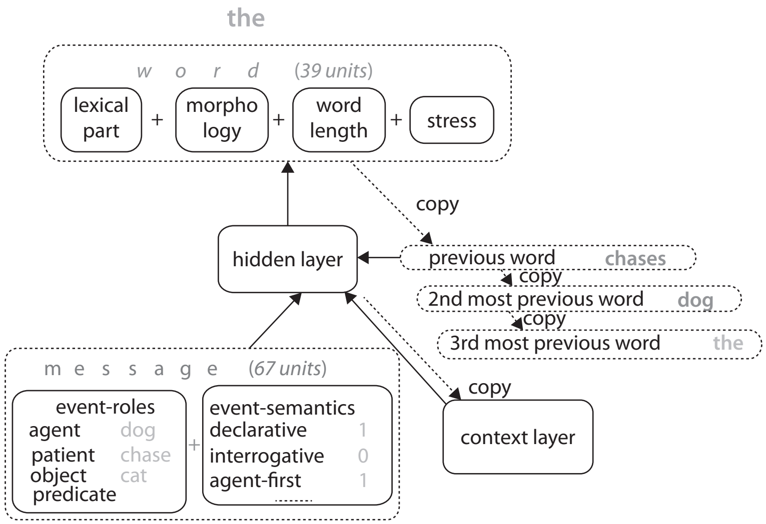

Figure 5. The architecture of the SP model. The figure refers to the English version and shows the production of the question 'Which dog chases the cats?' (graded pattern in output context layers = attenuated representations of less recent words)

The architecture was trained on an artificial grammar similar to that of the model of SC. Wh-questions were embedded in a larger set of sentence types that imposed a general bias for a SVO word order for both languages. Wh-Subject questions conformed to this bias, whereas wh-Object questions did not. The grammar was implemented in English and Modern Greek according to the characteristics of each language, again under the assumption that the frequencies of the sentence types applied to both (Redington & Thomas 2004).

The default version of the model was moderately successful in capturing data on wh-Question production in TD. Its main successes were simulating higher accuracy rates in wh-Subject questions than in wh-Object questions and the production of error patterns similar to those produced by TD children (e.g., gap-filling errors in English: 'What did the dog chase something?'; van der Lely & Battel 2003; semantic reversal errors in Modern Greek: 'Who chases the dog (cat)?'; Stavrakaki 2006). Its main shortcomings were that accuracy rates were too high for non-referential (who) Wh-object questions in the Modern Greek version and that the model produced a more restricted variety of error patterns some of which were not observed in children's production (e.g., single-noun responses, 'Which dog chases the dog?'). These shortcomings suggested the need for more detailed models of SP, which include mechanisms for variable-binding to prevent single-noun responses and allow for generalisation (see Chang et al. 2002), as well as language-specific measurements for sentence types and word orders.

With these successes and limitations in mind, we sought conditions of atypical processing that would generate profiles similar to those of children with SLI. The condition that was identified was the use of a low-discriminality activation function. This condition resulted in a drop in accuracy rates more pronounced in wh-Object questions, i.e., similar to the empirical data (van der Lely & Battel 2003; Stavrakaki 2006). In the English version of the model, this condition yielded responses like 'Which cat does the cat chases the dog?', i.e., sentences including a combination of errors (gap-filling and double-marking of tense). Thus, similar to the empirical data of van der Lely and Battel (2003), a type of ungrammatical response was produced exclusively in the SLI version. In the Modern Greek version, the atypical condition simulated the SLI data exhibiting the persistence of responses that are grammatically inacceptable (e.g., both nouns in accusative case; Stavrakaki 2006).

Discussion

The present study was motivated by the idea that models and theories of language acquisition should be general in two ways. First, they should address language acquisition in languages with different typological characteristics. Second, they should account for language acquisition in typical and atypical development. Based on this idea, we have designed models which each adopted a 2x2 research design, novel within the field of connectionist psycholinguistics. The first factor of the research design, cross-linguistic variation, considered two levels, English and Modern Greek. The second factor, typical-atypical development, considered TD and SLI. The research design was applied iteratively to three domains of language acquisition, namely IM, SC, and SP. This yielded three different neural network models, each having four versions.

This systematic modelling approach allowed for the explanation, under the connectionist framework, of a wide range of phenomena in the acquisition of the two languages of our cross-linguistic paradigm in TD and SLI. Data on SLI were addressed by implementing conditions of atypical processing constraints, corresponding to different aetiological considerations of the impairment, and evaluating the extent to which these produced developmental profiles symptomatic of the linguistic profile of SLI in the two languages. These findings were the focus of the current chapter. Considered together, they encourage the development of a unified explanation of the impairment, based on the modelling of the acquisition of three different linguistic domains in SLI, for two languages. This explanation would be indispensably tied to the way models accounted for corresponding empirical data on TD. This was due to our 2x2 research design, in particular due the fact that apart from the inclusion of atypical processing constraints to simulate SLI, the TD and the SLI version of each model were otherwise identical.

The implications of the current study for our understanding of SLI could therefore be based on two types of observation. The first type refers to properties shared by the three models, applying both to TD and SLI, cross-linguistically. This form corresponds to the way the neuroconstructivist approach was instantiated in our study, in the sense that it reflects a broader theoretical position on the key elements of the cognitive architecture underlying the acquisition of different domains of language, and the main features of the linguistic environment that are relevant to empirical effects. The second type of observation relies upon the combined results from the modelling of the cross-linguistic profile of SLI in three domains of language.

Common form of the three models

As the three models focused on different areas of linguistic knowledge, they considered different architectures, different training sets, and different types of input and output information. Two principles however, were common across the three models. The first principle was the multiple cues approach. The architecture of the three models had a common form, considering multiple types of information or cues presented on the input layer of the networks. For example, the model of IM was presented with information on the phonological form of a root, information on lexical semantics, grammatical class and targeted inflection. The SC model was presented with cues encoding the morphological structure of words, cues related to phonology and prosody, as well as information on the internal activation state of the network in the previous timestep.

The broader theoretical hypothesis under the multiple cues approach was that the learning of the different domains of language relies on the integration of the different types of information presented to the relevant mechanisms. To achieve ceiling levels of performance, the networks needed to discover the relevance of the information provided by different cues with respect to the tasks, and weight the different cues together to produce the desired response for all the mappings of the training set. The integration of different cues was flexible, in a number of ways. Some cues could be redundant (e.g., grammatical class in the learning of English IM); other cues could be more relevant to particular types of mappings (e.g., lexical semantics in the learning of irregular morphology); and other cues could be less or more important at different developmental stages (e.g., early stages of SC: reliance upon local cues; later stages: reliance upon global cues).

The second principle that was common across models was the presence of psycholinguistically-motivated frequency structures in the training sets of the models. These structures were based on estimates for the distribution of phonological, morphological, and syntactic characteristics of English and Modern Greek, derived from child-directed or adult language corpora. They were embedded into artificial languages, presented to the networks in a non-incremental fashion. Note, some models have presented training sets that start small and incrementally grow, under the assumption that the input to the child is initially limited, either in child-directed speech or via limitations to working memory systems. There is not a consensus that this assumption is the right one. We were neutral on this point, but used exposure on the full language environment to allow for the tractability of experimental design and the simplicity of the model, leaving out the need for extra simulations to balance out the effects of the initial composition of the training sets. This approach places a future burden on scaling up the models with more realistic corpora and potentially incremental training regimes so as to better correspond to data on child language development. For the purposes of this study, however, it was sufficient to demonstrate that the assumptions we included were sufficient to capture patterns of TD acquisition across languages.

The distributional characteristics of the input, interacting with the general learning mechanisms of the PDP systems, accounted for accuracy rates across different types of mappings, and error types exhibited by the English and the Modern Greek version of the models. For example, in the English Model of IM, omission errors were produced as a prototype effect of mappings which involved the production of the default form in the output layer of the network, whereas overgeneralisation errors were produced as a result of the phonological similarity of irregular mappings to regulars. In the Modern Greek version of this model, default forms were missing. However, a phenomenon parallel to the Optional Infinitive stage in English, concerning the overuse of 3rd singular forms in early stages of Greek language acquisition, was captured and accounted for on the basis of the high frequency of 3rd singular forms in the input.

These interactions between multiple-cue learning systems and the environment to which they were exposed yielded a starting point at which the same architectures could produce development of patterns characteristic of the acquisition of English or Modern Greek in TD when paired with appropriate training sets. It remained to be shown that the presence of atypical processing constraints could also lead to profiles symptomatic of SLI in the two languages.

Combining results from the simulation of SLI

The conditions of atypical learning constraints under which the linguistic profile of English and Greek children SLI was simulated were the use of fewer hidden units combined with less-salient phonological representations in the model of IM, the use of a lower recurrency factor in the connections from the hidden to the context layer in the model of SC, and the use of a lower temperature for the activation function in the SP model. Although the three conditions were not identical, they shared the correspondence to an account whereby general processing limitations are the underlying cause of the patterns of deficits presented by children with SLI. More widely, the results suggested that general processing limitations could provide a unified account of the impairment across different linguistic domains and across languages, with the manifestation of behavioural deficits in each domain and each language depending on an interaction between atypical processing constraints and the structure of the problem domain (Leonard 1998).

More specifically, the simulations targeting the linguistic profile of SLI had the following implications for theoretical accounts of the impairment:

1. In the model of IM, the profile of children with SLI could be simulated through a combination of two types of deficits, a general processing deficit, implemented using fewer units in the hidden layer of the model, and a perceptual deficit, implemented using weaker phonological representations (i.e., with reduced activations). Neither of the two conditions could meet the quantitative criteria for the linguistic profile of children with SLI on its own. The impaired version of the MIG suggested that it is possible that more than one type of deficit underlies the impairment. The linguistic profile of SLI could be a `double-hit' effect, i.e., the result of the interaction between the two different deficits. This account is, of course, not parsimonious unless one could demonstrate that the same developmental process that leads to weak phonological representations also leads to processing limitations in the IM system (for example, if those same processing limitations were to hamper the development of the phonological representations themselves). Nevertheless, the model demonstrated the mechanistic validity of the double-hit account to explain the data.
2. In each model, the same atypical constraint could account for the linguistic profile of children with SLI in Modern Greek. The same type of deficit altered the behaviour of the default model in a way that was consistent with the differences in the linguistic performance of children with SLI and TD children in each language. For example, the impaired version of the English model of IM simulated an increase in omission errors, while the impaired version of the Modern Greek model simulated the reliance of the perfective past tense on non-sigmatic forms. This points to accounts of the impairment positing

general processing limitations as the cause of SLI. These accounts could be general across languages, even though error patterns can differ.

3. The implementation of general processing deficits in our study addressed dissociations in the patterns of performance of TD children and children with SLI. It also accounted for the presence of a given dissociation in one language but not in another. In particular, in the model of SC, a drop in the rates in which passive sentences were (correctly) interpreted as patient-first was observed when atypical learning constraints were implemented in the Modern Greek version; however, a parallel deficit was not observed in the English version. A similar pattern was also present in the cross-linguistic empirical data (Dick et al. 2004; Stavrakaki 2002). The model suggested that the impairment in the interpretation of passives in the Modern Greek version was a result of the vulnerability of cues related to word-order when working memory resources were reduced. The deficit in passives was more pronounced in Modern Greek because of the use of case-marking in this language. The case-marking pattern of pasives sentences was a cue prompting the network to classify these sentences incorrectly, i.e., as agent-then-patient sentences. When cues related to word-order were attenuated, the interpretation of passives was more easily influenced by the case-marking cue.
4. It has been argued (van der Lely & Battell 2003) that unlike the behavioural data, general processing limitation accounts of SLI predict that the deficits of children with SLI in SP would involve mainly omissions of sentence constituents rather than responses involving commission errors, such as gap-filling errors ('What did the dog chase something?'). The English model of SP demonstrated that this criticism was unfounded. It simulated increased rates of gap-filling errors and accounted for these by demonstrating that under atypical processing constraints, the network tended to overapply the noun-verb-noun microstructure at greater rates.

Finally, the modelling work yielded testable predictions for the linguistic profile of SLI in Modern Greek, which could be examined in future empirical investigations. The strongest of these were as follows:

1. In tasks examining number agreement between the subject and the verb in the formation of the perfective past tense, children with SLI should produce overgeneralisations of the 3rd singular person of the past in other persons. These might be perfective or imperfective stems, and should occur more frequently than present 3rd singular forms (i-forms). Such errors should also be produced at higher rates in the 2nd person of the singular number, which present phonological overlap with the 3rd singular forms. These errors should also be observed in TD, although at earlier developmental stages and at lower rates.

2. In an online sentence interpretation task for Modern Greek considering the interplay between word-order information and local morphological cues (number and case-marking), children with SLI should rely more than TD children on morphological cues, because these are more salient than long-distance dependencies. In these tasks a positive effect of number agreement on accuracy rates should be pronounced mainly in O-Cl sentences ('It is the cat that the dogs chase') because these require this cue to be disambiguated. Case-marking would facilitate the interpretation of O-Cl sentences, and impede, to a lesser extent, the comprehension of passives. This interaction should be due to the fact that case-marking and word-order perform in coalition for the disambiguation of word-order in O-Cl sentences, and in competition in passives.

We have argued for the objective of unifying accounts of language acquisition across typical and atypical populations and across languages. We believe implemented computational models of language acquisition are valuable and complementary tools in this enterprise.

Acknowledgements

This work has been supported by the National Scholarship Foundation of Greece (IKY) and the Greek Ministry of Education. The research of the second author was supported by UK MRC Grant G0300188, the European Commission Grant NEST-029088 (ANALOGY) and UK ESRC grant RES-062-23-2721 and UK ESRC grant RES-062-23-2721. We would also like to thank Dr Martin Redington, for his help on the development of the second model described in this chapter.

References

Bishop, D. V. M. 1994. Grammatical errors in specific language impairment: Competence or performance limitations? *Applied Psycholinguistics* 15: 507–550. DOI: 10.1017/S0142716400006895

Bishop, D. V. M. 1997. *Uncommon Understanding: Development and Disorders of Language Comprehension in Children*. Hove: Psychology Press.

Chang, F., Dell, G. S., Bock, J. K. & Griffin, Z. M. 2000. Structural priming as implicit learning: A comparison of models of sentence production. *Journal of Psycholinguistic Research* 29: 217–229. DOI: 10.1023/A:1005101313330

Chang, F. 2002. Symbolically speaking: A connectionist model of sentence production. *Cognitive Science* 26: 609–651. DOI: 10.1207/s15516709cog2605_3

Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge MA: The MIT Press.

Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Chomsky, N. 1986. *Knowledge of language: Its Nature, Origin and Use*. New York NY: Praeger.

Chomsky, N. 1991. Some notes on the economy of derivation and representation. In *Principles and Parameters in Comparative Grammar*, R. Freidin (ed.), 417–454. Cambridge MA: The MIT Press.
Chomsky, N. 1995. *The Minimalist Program*. Cambridge MA: The MIT Press.
Chomsky, N. 1998. *Minimalist Iniquities: The framework*. Cambridge MA: The MIT Press.
Clahsen, H. & Dalalakis, J. 1999. Tense and agreement in Greek SLI: A case study. *Essex Research Reports in Linguistics* 24: 1–25.
Diamanti, M. 2000. Simple morphosyntactic structures in the speech of a child with Specific Language Impairment. In *Proceedings of the 8th Symposium of the Panhellenic Association of Logopedist*, 201–220. Athens: Ellinika Grammata. (In Greek)
Dick, F., Bates, E., Wulfeck, B., Aydellot, J., Donkers, N. & Gernsbacher, M. 2001. Language deficits, localization, and grammar: Evidence for a distributive model of language breakdown in aphasic patients and neurologically intact individuals. *Psychological Review* 108: 759–788. DOI: 10.1037/0033-295X.108.4.759
Dick, F., Wulfeck, B., Krupa-Kwiatkowski, M. & Bates, E. 2004. The development of complex sentence interpretation in typically developing children compared with children with specific language impairments or early unilateral focal lesions. *Developmental Science* 7(3): 360–377. DOI: 10.1111/j.1467-7687.2004.00353.x
Elman, J. L. 1990. Finding structure in time. *Cognitive Science* 14: 179–211. DOI: 10.1207/s15516709cog1402_1
Elman, J. L. 1991. Distributed representations, Simple Recurrent Networks, and grammatical structure. *Machine Learning* 7: 195–225.
Elman, J. L. 1993. Learning and development in neural networks: The importance of starting small. *Cognition* 48: 71–99. DOI: 10.1016/0010-0277(93)90058-4
Francis, W. & Kucera, H. 1982. *Frequency Analysis of English Usage*. Boston MA: Houghton Miffin.
Fromkin, V., Blair, D. & Collins, P. C. 2002. *An Introduction to Language*, 7th edn. Boston MA: Thomson Heinle.
Gathercole, S. E. & Baddeley, A. D. 1990. Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language* 29: 336–360. DOI: 10.1016/0749-596X(90)90004-J
Gopnik, M. 1990. Feature-blind grammar and dysphasia. *Nature* 344 (62–68): 715. DOI: 10.1038/344715a0
Gopnik, M. & Crago, M. 1991. Familial aggregation of the developmental language disorders. *Cognition* 39: 1–50. DOI: 10.1016/0010-0277(91)90058-C
Hatzigeorgiou, N., Spiliotopoulou, A., Vacalopoulou, A., Papakostopoulou, A., Piperidis, S., Gavriilidou, M. & Karayanis, G. 2000. Hellenic National Corpus (HNC): A modern Greek corpus on the internet. In *Proceedings of the 21st Annual Meeting of the Linguistics Department, School of Philology, Faculty of Philosophy, Aristotle University of Thessaloniki*, 812–821, May 2000, Thessaloniki. (In Greek)
Hoeffner, J. H. & McClelland, J. L. 1993. Can a perceptual processing deficit explain the impairment of inflectional morphology in developmental dysphasia? A computational investigation. In *Proceedings of the 25th Child Language Research Forum*, E. V. Clark (ed.), 38–49. Stanford CA: CSLI & Chicago IL: University of Chicago Press.
Holton, D., Mackridge, P. & Philippaki-Warburton, I. 2003. *Greek: An Essential Grammar of the Modern Language*. London: Routledge.

Hutzler, F, Ziegler, J. C., Perry, C., Wimmer, H. & Zorzi, M. 2004 Do current connectionist learning models account for reading development in different languages? *Cognition* 91(3): 273–296. DOI: 10.1016/j.cognition.2003.09.006

Joanisse, M. F. 2004. Specific language impairments in children: Phonology, semantics and the English past tense. *Current Directions in Psychological Science* 13(4): 156–160. DOI: 10.1111/j.0963-7214.2004.00297.x

Joanisse, M. F. & Seidenberg M. S. 1999. Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Sciences of the United States of America* 96: 7592–7597. DOI: 10.1073/pnas.96.13.7592

Joanisse, M. F. & Seidenberg, M. S. 2003. Phonology and syntax in specific language impairment: Evidence from a connectionist model. *Brain and Language* 86: 40–56. DOI: 10.1016/S0093-934X(02)00533-3

Kail, R. 1994. A method for studying the generalized slowing hypothesis in children with specific language impairment. *Journal of Speech and Hearing Research* 37: 418–421. DOI: 10.1044/jshr.3702.418

Kail, M. & Diakogiorgi, K. 1994. Morphology and word order in the processing of Greek sentences: A crosslinguistic and developmental perspective. In *Themes in Greek Linguistics, Current Issues in Linguistic Theory* [Current Issues in Linguistics Theory 117], I. Phillipaki-Warburton, K. Nicolaidis & M. Sifianou (eds), 325–332. Amsterdam: JoŠ Benjamins. DOI: 10.1075/cilt.117.48kai

Karaminis, T. 2012. *Connectionist Modelling of Morphosyntax in Typical and Atypical Development for English and Modern Greek*. PhD dissertation, University of London.

Karaminis, T. N. & Thomas M. S. C. 2012a. Connectionism. In *Encyclopedia of the Sciences of Learning*, Seel, Norbert M. (ed.), Part 3, 767–771. New York NY: Springer.

Karaminis, T. N. & Thomas, M. S. C. 2012b. Connectionist theories of learning. In *Encyclopedia of the Sciences of Learning*, Part 3, N. M. Seel (ed.), 771–774. New York NY: Springer.

Karaminis, T. N., & Thomas, M. S. C. 2010. A cross-linguistic model of the acquisition of inflectional morphology in English and Modern Greek. In *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, S. Ohlsson & R. Catrambone (eds), 730–735. Austin TX: Cognitive Science Society.

Leonard, L. B. 1989. Language learnability and specific language impairment in children. *Applied Psycholinguistics* 10: 179–202. DOI: 10.1017/S0142716400008511

Leonard, L. B. 1998. *Children with Specific Language Impairment*. Cambridge MA: The MIT Press.

Leonard, L. B., Bortolini, U., Caselli, M. C., McGregor, K. K. & Sabbadini, L. 1992. Morphological deficits in children with specific language impairment: The status of features in the underlying grammar. *Language Acquisition* 2: 151–179. DOI: 10.1207/s15327817la0202_2

Leonard, L. B., McGregor, K. K. & Allen, G. 1992. Grammatical morphology and speech perception in children with specific language impairment. *Journal of Speech and Hearing Research* 35: 1076–85. DOI: 10.1044/jshr.3505.1076

Lupyan, G. & Christiansen, M. H. 2002. Case, word order and language learnability: Insights from connectionist modeling. In *Proceedings of the 24th Annual Conference of the Cognitive Science Society*, 596–601. Mahwah NJ: Lawrence Erlbaum Associates.

MacWhinney, B. & Bates, E. 1989. *The Crosslinguistic Study of Sentence Processing*. Cambridge: CUP.

Mareschal, D., JoŠson, M. H., Sirois, S., Thomas, M. S. C. & Westermann, G. (eds). 2007a. *Neuroconstructivism, Vol. 1: How the Brain Constructs Cognition*. Oxford: OUP. DOI: 10.1093/acprof:oso/9780198529910.001.0001

Mareschal, D., Sirois, S., Westermann, G. & JoŠson, M.H. (eds). 2007b. *Neuroconstructivism, Vol. 2: Perspectives and Prospects*. Oxford: OUP.
DOI: 10.1093/acprof:oso/9780198529934.001.0001

Marshall, C.R. & van der Lely, H.K. 2006. A challenge to current models of past tense inflection: The impact of phonotactics. *Cognition* 100(2): 302–320.
DOI: 10.1016/j.cognition.2005.06.001

Marslen-Wilson, W.D. & Tyler, L.K. 1998. Rules, representations, and the English past tense. *Trends in Cognitive Sciences* 2: 428–435. DOI: 10.1016/S1364-6613(98)01239-X

Mastropavlou, M. 2003a. The Status of gender and agreement features in the grammar of two SLI Children. In *Proceedings of the 16th International Symposium of Theoretical and Applied Linguistics*, M. Mattheoudakis & A. Psaltou-Joysey (eds), 356–373, Thessaloniki: Aristotle University of Thessaloniki.

Mastropavlou, M. 2003b. Specific Language Impairment: Description of the speech of three children with SLI and the role of therapy in their development. In *Disorders of Communication and Speech: Prevention, Research, Intervention, and New TecŠologies in Health*, M. Glikas & G. Kalomiris (eds), 208–224. Athens: Ellinika Grammata.

Matthews, D.E. & Theakston, A.L. 2006. Errors of omission in English-speaking children's production of plurals and the past tense: The effects of frequency, phonology, and competition. *Cognitive Science* 30: 1027–1052. DOI: 10.1207/s15516709cog0000_66

Norris, D. 2005. How do computational models help us build better theories? In *Twenty-First Century Psycholinguistics: Four Cornerstones*, A. Cutler (ed.), 331–346. Mahwah NJ: Lawrence Erlbaum Associates.

Pinker, S. 1999. *Words and Rules*. London: Weidenfeld & Nicolson.

Plaut, D.C. 2000. Methodologies for the computer modeling of human cognitive processes. In *Handbook of Neuropsychology*, Vol. 1, F. Boller, J. Grafman & G. Rizzotti (eds). Oxford: Elsevier.

Plunkett, K. & Juola, P. 1999. A connectionist model of English past tense and plural morphology. *Cognitive Science* 23: 463–490. DOI: 10.1207/s15516709cog2304_4

Plunkett, K. & Marchman, M. 1991. U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition* 38(1): 43–102.
DOI: 10.1016/0010-0277(91)90022-V

Plunkett, K. & Marchman, V. 1993. From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition* 48(1): 21–69.
DOI: 10.1016/0010-0277(93)90057-3

Rice, M.L. 2000. Grammatical symptoms of specific language impairment. In *Speech and Language Impairments in Children: Causes, Characteristics, Intervention and Outcome*, D.V.M. Bishop & L.B. Leonard (eds), 17–34. Hove: Psychology Press.

Rice, M.L., Wexler, K. & Cleave, P. 1995. Specific Language Impairment as a period of extended optional infinitive. *Journal of Speech and Hearing Research* 38: 850–863.
DOI: 10.1044/jshr.3804.850

Richardson, F.M., Thomas, M.S.C. & Price, C.J. 2010. Neuronal activation for semantically reversible sentences. *Journal of Cognitive Neuroscience* 22 (5): 943–954.
DOI: 10.1162/jocn.2009.21238

Rumelhart, D.E., Hinton, G.E. & McClelland, J.L. 1986. A general framework for parallel distributed processing. In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations*, D.E. Rumelhart, J.L. McClelland & the PDP Group (eds), 45–76. Cambridge MA: The MIT Press.

Rumelhart, D. E., Hinton, G. & Williams, R. 1986. Learning internal representations by error propagation. In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations*, D. E. Rumelhart, J. L. McClelland & the PDP Group (eds), 318–362. Cambridge MA: The MIT Press.

Rumelhart, D. E. & McClelland, J. L. 1986. On learning the past tenses of English verbs. In *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 2: Psychological and Biological Models*, J. L. McClelland & D. E. Rumelhart (eds), 216–271. Cambridge MA: The MIT Press.

Rumelhart, D. E., McClelland, J. L. & the PDP Research Group (eds). 1986. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations*. Cambridge MA: The MIT Press.

Seidenberg, M. S. 2011. Reading in different writing systems: One architecture, multiple solutions. In *Dyslexia Across Languages: Orthography and the Gene-Brain-Behavior Link*, P. McCardle, J. Ren & O. Tzeng (eds), 146–148. Baltimore MD: Paul H. Brookes.

Smith, N. 2008. *Morphosyntactic Skills and Phonological Short-term Memory in Greek Preschool Children with Specific Language Impairment*. PhD dissertation, University of Reading.

Stamouli, S. 2000. The grammatical categories of tense, aspect, and subject-verb agreement: A comparative study of a Greek-speaking child with specific language impairment and a typical developing child. In *Proceedings of the 8th Symposium of the Panhellenic Association of Logopedists*, 159–174. Athens: Ellinika Grammata. (In Greek).

Stavrakaki, S. 2002. Sentence comprehension in Greek SLI children. In *Investigations in Clinical Linguistics and Phonetics*, N. Hewlett, L. Kelly & F. Windsor (eds), 57–72. Hillsdale NJ: Lawrence Erlbaum Associates.

Stavrakaki, S. 2006. Developmental perspectives on Specific Language Impairment: Evidence from the production of wh-questions by Greek SLI children over time. *Advances in Speech-Language Pathology* 8(4): 384–396. DOI: 10.1080/14417040600880714

Stavrakaki, S. & Clahsen, H. 2009. The perfective past tense in Greek child language. *Journal of Child Language* 36: 113–142. DOI: 10.1017/S0305000908008866

Stavrakaki, S., Koutsandreas, K., & Clahsen, H. 2012. The perfective past tense in Greek children with Specific Language Impairment. *Morphology* 22: 143–171. DOI: 10.1007/s11525-011-9194-5

Stephany, U. 1997. The acquisition of Greek. In *The Crosslinguistic Study of Language Acquisition*, Vol. 4, D. I. Slobin (ed.), 183–333. Hillsdale NJ: Lawrence Erlbaum Assoicates.

Tallal, P. & Piercy, M. 1973a. Defects of non-verbal auditory perception in children with developmental aphasia. *Nature* 241: 468–469. DOI: 10.1038/241468a0

Tallal, P. & Piercy, M. 1973b. Developmental aphasia: Impaired rate of nonverbal processing as a function of sensory modality. *Neuropsychologia* 11: 389–398. DOI: 10.1016/0028-3932(73)90025-0

Thomas, M. S. C. 2005. Characterising compensation. *Cortex* 41(3): 434–442. DOI: 10.1016/S0010-9452(08)70277-6

Thomas, M. S. C. & Karmiloff-Smith, A. 2002. Are developmental disorders like cases of adult brain damage? Implications from connectionist modelling. *Behavioral and Brain Sciences* 25: 727–788.

Thomas, M. S. C. & Karmiloff-Smith, A. 2003. Modeling language acquisition in atypical phenotypes. *Psychological Review* 110 (4): 647–682. DOI: 10.1037/0033-295X.110.4.647

Thomas, M. S. C. & Redington, M. 2004. Modelling atypical syntax processing. In *Proceedings of COLING 2004, Psycho-Computational Models of Human Language Acquisition,* W. G. Sakas (ed.), 87–94. Geneva: COLING.

Tsimpli, I. M. 2001. LF-interpretability and language development: A study of verbal and nominal features in Greek normally developing and SLI children. *Brain and Language* 77: 432–448. DOI: 10.1006/brln.2000.2413
Tzanidaki, D. I. 1995. Greek word order: Towards a new approach. *UCLWorking Paper in Linguistics* 7: 247–277.
Ullman, M. & Pierpont, E. 2005. Specific Language Impairment is not specific to language: The Procedural Deficit hypothesis. *Cortex* 41(3): 399–433. DOI: 10.1016/S0010-9452(08)70276-4
van der Lely, H. K. J. 1996. Empirical evidence for the modularity of language from Grammatical SLI children. In *Proceedings of the 20th Annual Boston University Conference on Language Development*, Vol. 2, A. Stringfellow, D. Cahana-Amitay, E. Hughes & A. Zukowski (eds), 781–791. Somerville MA: Cascadilla Press.
van der Lely, H. K. J. & Battell, J. 2003. Wh-movement in children with grammatical SLI: A test of the RDDR hypothesis. *Language* 79: 153–181. DOI: 10.1353/lan.2003.0089
van der Lely, H. K. J. & Ullman, M. 2001. Past tense morphology in specifically language impaired children and normally developing children. *Language and Cognitive Processes* 16: 177–217. DOI: 10.1080/01690960042000076