Constraints on language development

Insights from developmental disorders

Michael S. C. Thomas

1. Introduction

When one assesses the language abilities of children and adults with developmental disorders, it is not uncommon to find an uneven profile across the sub-domains of language. Standardized tests for various aspects of language can exhibit a differential relationship compared both to each other and to overall (average) mental age (MA). For example, in a comparison of Down syndrome (DS), Williams syndrome (WS), autism and Fragile X (FraX), Fowler (1998) described dissociations between phonology, lexical semantics, morphosyntax and pragmatics. From these dissociations, it is evident that general cognition cannot be a reliable indicator of all aspects of language function in children with learning disabilities. While language acquisition typically lags behind MA-level expectations in children with learning disabilities, Fowler noted that disorders such as Williams syndrome and hydrocephalus with associated myelomeningocele appear superficially to be exceptions. From her comparison, Fowler concluded that pragmatics and lexical semantics are more closely tied to MA than phonology and morphosyntax.

Tager-Flusberg and Sullivan (1997) carried out a similar comparison of the same four disorders but this time seeking possible asynchronies in the early development of semantic, grammatical, and pragmatic aspects of language. These authors also noted disparities in areas such as vocal development, social communicative development, gesture, lexical development, phonological development, early grammar and pragmatics.

However, despite the differences highlighted in their respective reviews, both Fowler (1998) and Tager-Flusberg and Sullivan (1997) also noted similarities across the disorders. For example in early development, there were

consistent patterns of errors displayed in speech articulation; and in morphosyntax, although some disorders stopped short of mastery, the order of acquisition of syntactic structures appeared similar. In some senses, atypical language development generally retains some link with the profile of normal development.

What can this pattern of commonalities and dissociations tell us about the development of the language system? Two explanatory frameworks compete to interpret the results. One approach is based on the assumption of functional modularity in the normal adult system. The field of neuropsychology has identified case studies of healthy adults who exhibit selective deficits to different components of language following acquired brain damage. From these dissociations, a modular functional architecture has been inferred. Within modular theories, the linguistic performance of individuals with developmental language impairments is viewed as reflecting the architecture of the normal system but with selective components of this system under-developed or over-developed (Clahsen & Temple 2003). This framework provides a comfortable fit between the results of standardized language tests and atypical functional structure. Assuming we have tests that index the integrity of individual modules (e.g., tests of vocabulary, tests of grammar, tests of phonological awareness, and so on), scores in the normal range can be read off as reflecting a normally developed component and scores above or below the normal range can be read off as reflecting an (atypically) over- or under-developed component. This mapping of test results to modular structure in *developmental* disorders rests on one of two assumptions. Either the modular system identified in the adult is also present in the infant, so that language development can commence with an initial selective anomaly in one or more components; or the modular structure emerges through development in such a way that when things go wrong, some parts emerge with atypical functionality while the rest nevertheless manage to emerge displaying their normal functionality. Together, these alternatives constitute the assumption of *residual normality* (Thomas & Karmiloff-Smith 2002a). One further assumption is required for us to read off a normal score achieved on a standardized test as a guarantee of the normal functioning of an underlying component: that atypical cognitive processes could not generate the same normal score on this test.

The alternative framework, sometimes referred to as neuroconstructivism (Karmiloff-Smith 1998), places a much greater emphasis on the role of development in producing cognitive structure. It is based on the premise that the adult modular structure is not present in the infant but is itself a product of the developmental process. This is a view strongly motivated by data from de-

velopmental cognitive neuroscience (Elman, Bates, Johnson, Karmiloff-Smith, Parisi, & Plunkett 1996; Karmiloff-Smith 1998). This developmental perspective draws into question the sensitivity of standardized tests, raising the possibility that scores in the normal range may be achieved by atypical cognitive processes. Instead it is argued that sensitive on-line tasks are necessary to properly assess underlying processes (Karmiloff-Smith 1997; Karmiloff-Smith, Thomas, Annaz, Humphreys, Ewing, Grice, Brace, Van Duuren, Pike, & Campbell in press). In the view of these authors, the clean pattern of normal versus impaired modules identified in some developmental disorders may in part be an artifact of the straightjacket of standardized tests. If a child takes a receptive vocabulary test, they can only possibly score below, at or above the normal range.

The debate between these two explanations of uneven linguistic profiles has at times become polarized. On the one hand, there are strong claims that for given developmental disorders, certain cognitive structures *must have* developed normally given behavior in the normal range (sometimes these are referred to as 'intact' or 'spared' systems). On the other hand there are counter claims that since the developmental processes we know about could not have produced such an uneven modular outcome, the relevant behavior *must be* produced by structures that are qualitatively different and atypical. For example, such polarization has occurred in evaluating syntax processing in Williams syndrome, and in evaluating the lexicon in the so-called 'grammatical' subtype of Specific Language Impairment.

Although my own previous work has been carried out within the neuroconstructivist framework, in this chapter my intention is to step back from this debate somewhat, and focus on exploring the notion of *constrained development*. This is because both frameworks must eventually incorporate an account of this sort, even if the strength of the constraints will differ in the two types of account. In the next section, I consider how both modular and neuroconstructivist frameworks still face significant challenges in characterizing the developmental process.

2. Development produces the disorder

In an older child, adolescent, or adult with a developmental language disorder, development has played some role in producing the observed behavioral deficits. The exact contribution of development is disputed. However, in both modular and neuroconstructivist frameworks, the nature of the developmental process remains obscure.

The modular approach de-emphasizes the contribution of development, placing the antecedents of deficits in particular components of a proto-language system already present in the infant. For example, various explanations of Specific Language Impairment (SLI) exist which propose a deficit restricted to abstract language structures involved in the rule-governed movements or combinations of words into complex structures (see Ullman & Pierpont in press, for review). According to different versions, children may come to language impaired in their ability to establish structural relationships in sentences, such as agreement or specifier head-relations; or they may lack rules for linguistic features; or they may be stuck in a period of language development where marking of tense is taken to be optional; or they may be solely impaired on non-local dependency relations; or they may have problems with more general language functions such as learning implicit rules. The implication in a disorder argued to have a strong genetic component is that such impairments pre-date acquisition.

Two aspects remain vague in the modular account. The first is the exact granularity of the proto-language system, that is, the miniature, content-free modular functional architecture present in the pre-linguistic infant (see Thomas & Karmiloff-Smith in press, for discussion). The second is the developmental process by which this architecture acquires its content when exposed to a given social and language environment. While some researchers argue there is scant empirical evidence for the existence of adult-like modular functional structure in the infant (e.g., Bates & Roe 2001; Elman et al. 1996), here we need merely point out that if one is going to argue for such a structure, one needs to say exactly what it looks like. At what level of detail do functional distinctions exist in the infant system – between sounds, meanings, motor actions, and social interactions; or between phonology, morphology, syntax, and the lexicon? Stipulating the granularity of the infant proto-system permits specification of which components will have the initial developmental deficit. The account must then be complemented by specification of the processes of learning. Such processes must put particular content in each of the modular 'boxes' whilst allowing the components to interact fluidly in language comprehension and production. Even a strictly modular account of atypical language development needs to postulate firstly a startstate (however much constrained) and secondly a pathway via a set of interactions with an information-laden world to arrive at the final uneven structures observed in the adult developmental disorder.

While the neuroconstructivist approach accepts functional modularity as a possible characterization of the adult system, it rejects it as a startstate for

the infant cognitive system. This approach rests on a theory that modularity emerges as a product of development, from a relatively less differentiated information processing system. The less differentiated system has capacities that are *relevant* to cognitive domains rather than specific to them (for instance, ability to processes sequences may be relevant to syntax processing, without being specific to the linguistic structures that sentences contain). This initial flexibility is lost across development as the system commits its relevant capacities to particular domains. An explanation of developmental deficits consists in identifying how these initial domain relevancies have been altered in the disorder, and then how the subsequent process of emergent modularization has been perturbed (if indeed it has been). An emphasis on differences in the startstate leads neuroconstructivists to investigate the infant precursors of later uneven cognitive profiles (Karmiloff-Smith 1998). For example, Paterson, Brown, Gsödl, Johnson and Karmiloff-Smith (1999) noted that in adults with WS and DS, individuals with WS were relatively stronger than those with DS in language but the reverse was the case in the domain of number. When Paterson et al. explored the precursors of these cognitive skills in infants with the disorders, they found no advantage for toddlers with WS over DS in a language task, and *better* performance in WS than DS in the number task. The adult pattern was not replicated in the infant state, implying that different atypical developmental trajectories separate the populations (see Singer Harris, Bellugi, Bates, Jones, & Rossen 1997; Mervis & Robinson 2000, for discussion). These authors therefore argued against the atypical infant proto-cognitive system containing a miniature version of the adult functional structure with the same pattern of strengths and weaknesses.

Two difficulties remain for the neuroconstructivist approach. The first difficulty is not unrelated to the one faced by the modular approach. Even if there are much weaker constraints on the startstate of the proto-language system, these still need to be identified. What is the set of initial domain-relevancies that pre-date language, and what is the nature of the process that eventually delivers domain-specific functional structures? The account eventually needs to be concrete enough to establish the strength of the constraints governing the emergence of modularity; what "seeds" the proto-language systems starts with; what conditions would be sufficient to disrupt it; and how a genuinely "atypical" functional structure would behave. Presumably, even a system dealing in no more than domain relevancies must arrive with possible channels of information flow established – for example, between motor systems driving articulation, perceptual systems interpreting input, multi-modal systems linking to conceptual knowledge, and pragmatic systems linking with social

and emotional systems. The second difficulty is that while neuroconstructivism prompts its adherents to build developmental trajectories from infancy through childhood to adult language structures, the empirical basis for "proto" cognitive structures is problematic. For example, Paterson et al. (1999) compared scores on receptive vocabulary tests in adults (i.e., selecting a picture that goes with a word from a set of alternatives) with performance on a preferential looking task in infants, where infants were presented with two pictures (e.g., a dog and a cat) and their gaze behavior monitored when they heard a label (e.g., "Look at the dog! Look at the dog!"). Where differences are found in the adult and infant pattern in a cross-syndrome comparison, how do we know that the two 'vocabulary' tasks are indexing the same mechanism? The problem even hold when the same task is used – how can one be sure that the same task is treated the same way at very different ages? We need to be able to rule out the possibility that data showing differential relative profiles in infancy and adulthood not in fact the results of measuring different cognitive capacities at the two ages (such as, in the preceding example, lexical knowledge in the adult and attention/degree of novelty preference in the infant).

It is worth pointing out that evidence from brain imaging studies is often introduced in an attempt to distinguish modular and neuroconstructivist positions (predominantly by the latter group, as you may guess from the 'neuro' prefix). I won't discuss brain level evidence here, other than to suggest that it indicates that the effects of genetic mutations on brain development in developmental disorders tend to be widespread rather than focal (see Mareschal, Johnson, Sirios, Spratling, Thomas, & Westermann forthcoming; Karmiloff-Smith & Thomas 2003; Thomas 2003, for discussion); and that brain evidence has been interpreted both within modular and neuroconstructivist frameworks (e.g., for WS, see Reiss, Eckert, Rose, Karchemskiy, Kesler, Chang, Reynolds, Kwon, & Galaburda 2004; for a more modular perspective; and for a more neuroconstructivist perspective, Grice, Spratling, Karmiloff-Smith, Halit, Csibra, de Haan, & Johnson 2001; Karmiloff-Smith 1998; Neville, Mills, & Bellugi 1994; Mills, Alvarez, St. George, Appelbaum, Bellugi, & Neville 2000). Brain evidence remains problematic in that while it is suggestive, for instance in the lower degree of functional localization and specialization observed in the infant neocortex (Karmiloff-Smith 1998), it is not clear how brain function constrains the cognitive structures it is supporting at an given point in time (see Mareschal et al. forthcoming, for discussion).

Thus far, then, we have suggested that explanations of uneven language profiles are compromised by lack of an explicit developmental account of the origin of the architecture of the adult system. In the current chapter, I address

this issue as follows. First I characterize some of the properties a developmental account should have with reference to the multiple components of the language system. Second, taking the example of Williams syndrome, I indicate the type of empirical evidence that might be used to identify the (atypical) constraints operating on development in a disorder. Third, I discuss some recent findings from computational modeling, a forum that permits a more precise exploration of the way in which atypical constraints on development could produce behavioral deficits in a given language domain.

3. Characterizing the developmental process

A cognitive-level developmental theory that explains the uneven language profile found in some disorders must emphasize three characteristics: *interactivity, compensation*, and *timing* (Thomas & Karmiloff-Smith in press). In this section, I concentrate in the main on the first two of these (see Elman et al. 1996, for a more detailed consideration of timing).

3.1 Interactivity

Several authors have argued that early language development is characterized by *interactions* between multiple sources of information and the components that process them (e.g., Bishop 1997; Chiat 2001; Karmiloff-Smith 1997, 1998; McDonald 1997). For example, Chiat (2001) maintained that language acquisition should be construed as a mapping task between sound and meaning, through which the words and sentence structures of a language are established. To achieve this mapping, multiple sets of information are exploited. When semantics is ambiguous, phonology can be used to bootstrap the extraction of meaning. When phonology is ambiguous (for instance during lexical segmentation), semantics can be used to bootstrap the extraction of word-sound information. Together, phonological and semantic information help bootstrap the acquisition of morpho-syntax. In a developmental disorder where there are indications of differential deficits across the components of the language system, any explanation of behavioral impairments must incorporate the altered pattern of interactions (and their timing). Chiat (2001) carried out this exercise for SLI and favored an account that considers the language deficits in morphology and syntax as arising from impaired phonological processing. The phonological impairment then leads to consequent disruption of the interactions inherent in the mapping process.

A phonological account of SLI is consistent with the view that higher-level language deficits arise as a developmental consequence of lower level deficits, so that, for instance, the phonological impairments in SLI may themselves originate in low-level auditory processing problems. However, this theory is controversial in as much as some adults with SLI do not demonstrate low-level processing deficits in auditory discrimination (McArthur & Bishop 2004; Rosen 2003). One response is to postulate that auditory processing impairments may exist early in development and yet fail to be measurable in the mature system. This would be one instantiation of the claim that *timing* is an essential factor in producing developmental impairments. While the failure to find an auditory processing deficit in an adult with SLI cannot be assumed to mean that such a processing deficit did not exist in infancy and make an impact on early language development, the postulation of unmeasurable causal factors is problematic. One might argue that falsifiability of the low-level deficit theory is compromised if one assumes that the source of an adult language problem lies in a cause that can no longer be measured. Of course, this simply highlights the point that developmental deficits demand that empirical data are collected across the course of development rather than just at its endpoint. The early deficit theory is eminently testable using longitudinal studies in children with SLI or infants at risk for SLI.

The idea that low-level auditory processing deficits explain higher-level language problems in SLI is not supported as a *sufficient* condition by data comparing children with SLI and those with mild hearing impairments. Norbury, Bishop and Briscoe (2001) discovered phonological processing problems in both group but problems in productive inflectional morphology only in the SLI group. It appears that poor auditory processing is not necessarily associated with deficits in the more abstract, high-level aspects of language. In addition, even accepting the role of phonology, the causal pathway linking problems at this level to circumscribed syntactic difficulties (e.g., subject-verb number agreement) is at best obscure (though see Joanisse 2000, for some preliminary attempts to make these links in the domain of anaphor resolution).

Nevertheless, at a broad level, the importance of the quality of language input has been emphasized by a comparative analysis carried out by McDonald (1997), which contrasted several typical and atypical populations that exhibited either successful or unsuccessful acquisition of language. These populations included late L2 learners, deaf sign-language learners, individuals with Down syndrome, individuals with Williams syndrome and children with SLI. McDonald concluded that good representations of speech sounds were key in predicting the successful acquisition of a language including its syntax, again

supporting the view that the components of the language system interact across development.

3.2 Compensation

The second characteristic that any theory of atypical language development must incorporate is *compensation*. The importance of compensation can be illustrated by a triangular comparison of adult aphasics, healthy children who have experienced early focal brain damage, and children with developmental disorders (see Karmiloff-Smith & Thomas 2003; Thomas 2003). The comparison goes as follows. (1) Following focal brain damage to their left hemispheres, adults can show persistent selective deficits in their language abilities (e.g., as exhibited in non-fluent and fluent aphasia). However, (2) following similar focal damage, healthy children usually then go on to demonstrate recovery from initial aphasic symptoms and later perform within the normal range on language tasks (see Bates & Roe 2001, for a review). Presumably, the greater effective plasticity of the child brain has permitted compensation and reorganization of function. As a consequence, when we (3) compare adults who had focal lesions when they were children with adults who have developmental disorders of language, we find significant deficits only in the latter. Of course, pointing to the presence of deficits in a developmental disorder is somewhat tautological, but the comparison nevertheless raises the question that if genetic developmental disorders of language are to be characterized by initial selective deficits to language-relevant structures, why has compensation-to-recovery not occurred as it does in the children with early focal lesions? The answer is that compensation in the developmental disorder probably *has* occurred, but the constraints of the system are insufficient to allow performance to develop to a level within the normal range (Mareschal et al. forthcoming; Thomas 2003). This must be true for behaviorally defined disorders, because any child that had successfully compensated for their initial deficit would not be diagnosed as having a disorder. There are parallels to be drawn between healthy children with early acquired brain damage and those with developmental disorders, but the relevant comparison is for healthy children who have experienced widespread and/or diffuse brain damage rather than focal lesions (Thomas 2003).

Our account of the emergence of differentiated language structure in the adult will therefore need to incorporate interactivity, compensation, and timing, whether the early infant system is strongly or weakly constrained. This has significant implications for uneven profiles found in developmental disorders.

If we propose that the uneven profile can be explained by an initial deficit to a single component of the system (say, the proto-phonological system, proto-lexicon, proto-syntax system, or proto-pragmatic system), why wouldn't this impairment become smeared across other components through the interactions that occur between them during development? And why wouldn't other components in the system manage compensate for this selective deficit and so attenuate the impairment across development?

To take interactivity, if there were an initial selective impairment in pragmatics in infants with autism, one might expect the deficit to be passed to the lexicon, where words or phrases whose meaning can only be inferred from speaker intentions should not be acquired normally. One might expect non-canonical syntactic constructions (such as passives or cleft constructions) to be poorly processed, since these are predominantly employed in service of emphasizing the topic of the sentence for the listener, that is, for pragmatic reasons. To take compensation, if there were an initial selective impairment in syntax in SLI, why shouldn't the child compensate by using the lexicon to acquire common whole inflected forms and syntactic phrases, to be deployed in the appropriate communicative context and so avoiding diagnosis as having a language impairment? The exact answers to these questions are not important in the current context (perhaps both phenomena occur; see Section 5 for further discussion of SLI). The point is that uneven language profiles may encourage the idea that selective damage has occurred, but explanations must be couched in terms of the development of differentiated language structures. If theories propose highly selective deficits in the adult with the disorder, then they must incorporate *developmental* reasons why neither interactivity nor compensation has taken place.

If one is to build an explanation of language deficits in terms of the developmental process, what type of empirical evidence should guide one's hand? In this next section, I use Williams syndrome as an illustration.

4. The example of Williams syndrome

Williams syndrome involves the deletion of some 25 genes from one of the copies of chromosome 7 (see Donnai & Karmiloff-Smith 2000, for full details of the syndrome). Individuals with WS usually present with IQs in the 50–60s range, with poor spatial and numerical cognition. While there is an initial delay in language development, by adolescence and adulthood many individuals display large vocabularies that co-exist with relatively good scores on standard-

ized grammatical tests. Their language can include rich syntactic structure, with production and comprehension performance on complex syntactic structures (passives, relatives) in line with MA controls (Clahsen & Almazan 1998; Zukowski 2001).

In some respects, the developmental trajectory for language appears normal in WS. Thus, Mervis, Morris, Bertrand and Robinson (1999) noted that, while the syntactic abilities of children with WS (39 children from 2 years 6 months to 12 years of age) were considerably delayed, syntactic complexity was nonetheless appropriate for the mean length of utterance (MLU). This contrasts with DS, autism and FraX, where syntactic complexity turned out to be less than would be expected at MLUs over 3. This result prompted Mervis et al. to claim that WS is the first syndrome in which the normal relation between utterance length and complexity has been demonstrated. However, in other respects, the pattern is atypical. There are more errors in morphology (verb tense agreement, personal pronouns, grammatical gender; Karmiloff-Smith et al. 1997; Volterra, Capirci, Pezzini, Sabbadini, & Vicari 1996) than in syntax. Mervis et al. (1999) found that while the syntactic complexity scores of children with WS were significantly higher than would have been expected on the basis of spatial constructive ability, they were nevertheless significantly lower than would have been expected on the basis of receptive vocabulary ability, verbal ability, or auditory short-term memory. Across a large sample of 77 individuals between 5 and 52 years, Mervis et al. (1999) reported that performance on the Test of Receptive Grammar (Bishop 1983) was poor for complex constructions. Only 18% of the participants (22% of the adults) passed the test block that assessed relative clauses and only 5% (9% of the adults) passed the block assessing embedded sentences.

Such fractionation – patterns of strengths and weaknesses – appears in other areas of the WS language system (Thomas in press a). Pragmatics, less advanced in WS than grammar, also exhibits within-domain fractionation. There is relatively good performance in social sensitivity (e.g., making dyadic eye contact, sensitivity to non-verbal cues) but problems in areas such as greeting behaviors, topic maintenance, and question answering (Semel & Rosner 2003). In lexical-semantics, a relative strength in category concepts (e.g., the distinction between animals. Tools, clothing, furniture etc.) contrasts with problems understanding semantic relational concepts such as spatial-temporal terms (Phillips, Jarrold, Baddeley, Grant, & Karmiloff-Smith 2004). Even within category concepts, recent evidence has indicated differential naming problems across categories (Temple, Almazan, & Sherwood 2002; Thomas & Redington 2004), and it has been argued that the lexicon is an area of specific anomalies

in WS (Clahsen & Almazan 1998; Rossen, Klima, Bellugi, Bihrle, & Jones 1996; Temple et al. 2002).

In order to consider the developmental origins of this uneven pattern, researchers have turned to precursors of language in WS infants. In Karmiloff-Smith and Thomas (2003), we recently reviewed this work. The most salient aspect of the onset of language in WS is that it is delayed. Although this delay is variable, one study of 54 children with WS found an average delay of 2 years, similar to that found for children with Down syndrome (DS) (Singer Harris et al. 1997; see also Paterson et al. 1999). Though delayed, some aspects of early development reveal normal behavioural patterns. For example, the onset of hand banging predicts the onset of canonical babbling in infants with WS in the same way as it does in typically developing infants (Masataka 2001; Mervis & Bertrand 1997).

Despite the fact that phonological memory appears as a relative strength in WS in childhood and adulthood (Mervis et al. 1999), a study of the ability of infants and toddlers with WS to segment the fluent speech stream into words revealed serious delays (Nazzi, Paterson, & Karmiloff-Smith 2003). In part, then, language delays may be due to problems with the early development of speech perception and phonological representations.

However, some precursors appear not just delayed but atypical. For example, Laing and colleagues examined socio-interactive precursors to language development in toddlers with WS compared with MA controls (Laing, Butterworth, Ansari, Gsödl, Longhi, Panagiotaki, Paterson, & Karmiloff-Smith 2002). Although toddlers with WS were proficient at dyadic interactions with a caregiver (and indeed sometimes exceeded the scores of MA controls due to persistent fixation on the caregiver's face; see also Bertrand, Mervis, Rice, & Adamson 1993; Jones, Bellugi, Lai, Chiles, Reilly, Lincoln, & Adolphs 2000), there was a marked deficiency in triadic interactions incorporating an object. Specifically, toddlers with WS had difficulty switching attention from the caregiver to an object that was being referred to in communication (via pointing, looking, and naming). Such a deficiency could disadvantage the toddlers with WS in learning the names of objects, since shared attention to newly named objects is one of the main routes into vocabulary acquisition. And indeed, there is accumulating evidence that precursors to vocabulary development in WS are atypical.

Typically developing infants use the presence of linguistic or gestural information that accompanies the introduction of novel objects to influence their subsequent categorisation of those objects, sometimes over and above the perceptual similarities among the objects. However, Nazzi and Karmiloff-Smith

(2002) found that 2- to 6-year-old children with WS were significantly less able than typical controls to use verbal cues to constrain categorisation. Masataka (2000) found a similar poverty in the ability of 2–3 year olds with WS to use gestural information to constrain categorisation.

In typically developing children, the ability to use pointing to refer to objects tends to emerge before the use of verbal labels for the same purpose. Presumably, pointing indexes the emergence of the cognitive ability to make reference, prior to the lexical manifestation. Pointing to objects and eliciting pointing behaviour in adults also facilitate the ability to find the correct referent for a given label. However, in WS, Mervis and Bertrand (1997) found that the order was reversed, with the onset of productive vocabulary *preceding* pointing. Laing et al. (2002) confirmed a deficit in the pointing behaviour of infants with WS, despite relative proficiency at fine motor skills. Vocabulary acquisition, therefore, appears to rely on a different set of cues and constraints in WS. When Stevens and Karmiloff-Smith (1997) examined the constraints that older children and young adults with WS were using to learn novel words, these, too, appeared atypical.

Relations between markers of semantic knowledge and productive vocabulary were also unusual in young children with WS. Spontaneous exhaustive sorting of objects (such as arranging toy animals and blocks into their separate categories) indexes the development of semantic knowledge and tends to precede a rapid rise in the rate of vocabulary acquisition in typically developing children. By the time children find it clear which categories objects fall into, it becomes increasingly easier for them to attach consistent labels to different objects. However, for children with WS, Mervis and Bertrand (1997) found no evidence that exhaustive sorting preceded the vocabulary spurt. Indeed, several children with WS exhibited the reverse pattern – unlike children with DS who always displayed the normal pattern.

Finally, there is preliminary evidence that compared to normal children the vocabulary of young children with WS exhibits a reduced advantage for comprehension vocabulary over production vocabulary (Paterson 2000), implying a relatively higher productive vocabulary for their level for comprehension.

In sum, the study of precursors to language development in WS reveals two main themes. First there is an overall delay, perhaps of a more generalized nature incorporating delays in at least motor, phonological, and semantic development. Second, when language development gets underway, a differential balance emerges between the ability to encode and produce word forms on the one hand, and the acquisition of the semantic underpinnings for those words on the other. However, characterization of the endstate language sys-

tem in WS found in adolescents and adults remains controversial. Thomas and Karmiloff-Smith (2003) recently identified two main schools of theory. The first of these is more or less a null hypothesis. The *Conservative hypothesis* argues that the language we see in WS is not markedly atypical, just the product of delayed development combined with low IQ. The second school of theory develops the themes emerging from the study of early WS language development: the *Semantics-Phonology Imbalance hypothesis* comprises a cluster of claims that the WS language system involves a differential pattern of impairments across language.

The Conservative hypothesis runs as follows. Deficits in syntax and pragmatics in WS are what one might expect at a given level of mental retardation. Language development from the earliest age reflects the interests of a child with WS, specifically a strong desire for social interaction (e.g., Jones et al. 2000). Language is initially used more to mediate these interactions than as a referential tool. Subsequent vocabulary development reflects the special interests of the child with some degree of mental retardation, with unusual ('precocious') word usage employed as a strategic device to gain attention and mediate social interaction (Thomas & Redington 2004). Deficits that do exist in vocabulary reflect other non-linguistic aspects of WS. For instance their visuo-spatial processing deficit leads to problems acquiring spatial vocabulary (Phillips, Jarrold, Baddeley, Grant, & Karmiloff-Smith 2004). The challenge for the Conservative hypothesis, however, is to explain why individuals with WS should show errors in, for instance, morphosyntax, that are not found in typically developing children, and why they should show predominantly successful language acquisition when individuals with other genetic syndromes involving mental retardation do not. To the latter point, one could respond that it is the other disorders that have the problems (say, in phonology, while in WS, after a delay, this develops within the normal range). Tager-Flusberg, Plesa-Skwerer, Faja and Joseph (2003: 10) provide a recent statement of the Conservative position: "Despite claims to the contrary... there is no evidence that children with WS acquire language any differently than other [typically developing] children, although they may be delayed in the onset of first words and phrases, as would be expected given their mental retardation."

By contrast, the Semantics-Phonology Imbalance hypothesis (really a cluster of related hypotheses) argues that language development in WS takes place under altered constraints. Several atypical constraints have been proposed. First, there is the idea that individuals with WS have a *particular strength in, or a sensitivity of, phonological short-term memory* (Majerus 2004; Majerus, Palmisano, van der Linden, Barisnikov, & Poncelet 2001; Mervis et al. 1999).

For example, Vicari, Carlesimo, Brizzolara and Pezzini (1996) have labeled language in WS as "hyper-phonological", and Bishop (1999) has argued that WS demonstrates the importance of short-term memory for speech sounds in determining the success of language development. Second, there is the proposal the WS exhibits *a particular weakness in lexical semantics*. Volterra and colleagues have noted that grammatical problems in WS are especially evident with those aspects of morphology carrying out a semantic function; and that individuals with WS perform better than mental-age match controls only in those areas of language where semantic aspects are not involved (e.g., Pezzini, Vicari, Volterra, Milani, & Ossella 1999; Volterra, Capirci, & Caselli 2001). Rossen et al. (1996) proposed that anomalous activation dynamics within the lexicon, specifically impaired inhibitory dynamics mediating context effects, lead to imprecise knowledge of concepts in WS and atypical vocabulary usage (see Temple et al. 2002, for a similar proposal; Thomas, Docrell, Messer, Parmigani, Ansari, Karmiloff-Smith submitted, for discussion). Third, there might be *a lag between the development of phonology and semantics* in WS, or a *problem integrating the two sources of information*. For example, Karmiloff-Smith, Tyler, Voice, Sims, Udwin, Howlin and Davies (1998) found that when individuals with WS monitored a sentence for a target word, performance was like controls in showing disruptions following syntactic violations, but there was a divergence when those violations involved lexically based information. Here the control group showed disruption of word monitoring, but the WS group did not. This led the authors to propose that in WS, there is a deficit in integrating lexical-semantic information with phonological information in real-time processing. Indeed, Frawley (2002) subsequently argued that WS language should be seen primarily as a disorder involving integration deficits between processing modules.

In all of these cases, the outcome of the imbalance is a system that relies (or has relied at certain points in its developmental history) more on phonological information than semantic information, with certain consequent behavioral impairments. A complication of the Imbalance theory is that most of its components are logically independent and not mutually exclusive. It is at least possible that several of the hypotheses could conjointly turn out to be true. For example, WS might constitute a case where there are differences in phonology and in semantics, in a system exhibiting general delay and overlying effects of mental retardation.

From the example of Williams syndrome, then, we can see an initial characterization of an uneven language profile in adolescence and adulthood, including claims that grammar has (selectively) developed normally (Clahsen &

Almazan 1998). However, this initial modular proposal was not accompanied by any proposals for the developmental pathway (Thomas & Karmiloff-Smith in press). Moreover, the characterization was tempered by the fact that even syntax development is delayed in WS and then does not reach normal levels of mastery. Subsequent testing has revealed a good deal of fractionation or unevenness of different aspects of WS language, including within syntax, the lexicon, and pragmatics. This raises questions of whether a modular account of WS language could possibly deal with the granularity of fractionation by postulating one or more deficits to the initial proto-language system. To do so would seem to require implausible levels of detailed structure in the infant pre-linguistic system. The search for a developmental account then led to a focus on infant precursors, and here there accumulated evidence that some precursors to language were themselves atypical, for instance the deficit in triadic but not dyadic interaction, and the markers of referential communication. Although some researchers still prefer a "delayed but normal" explanation of WS language development, there is now a cluster of accounts that view this process in terms of an atypical balance between the lexical-semantic and phonological constraints, the former relatively weaker and the latter relatively stronger. In these accounts, the relatively high level of syntactic performance would be associated with the basis of good (albeit delayed) phonology. However, discussions still persist concerning whether syntax development itself follows a 'normal' course, and if it does, what this tells us of the constraints guiding typical and atypical language acquisition.

5. Computational investigations into constrained development

The methodology of computational modeling forms a convergent approach to understanding constraints on development, and how atypical constraints may produce sub-optimal development. Computational models provides a concrete basis to investigate more precisely how sources of information interact in the acquisition of a particular language domain, including opportunities for compensation, and the different ways in which delay and deviation may emerge from a system learning a facet of language. As with all methodologies, there are some limitations. Modeling necessarily involves simplification, and thus far it has focused in the main on individual domains (lexical segmentation, vocabulary acquisition, inflectional morphology, syntax processing; see Christiansen & Chater 2001) rather than the development and operation of multi-component systems (see Thomas & Karmiloff-Smith 2002a; Thomas &

Richardson in press, for discussion). Nevertheless, work to date has generated insights into the potential causes of developmental language deficits.

One of the main modeling formats applied to developmental disorders has been that of connectionist networks (see Thomas & Karmiloff-Smith 2002b, for a review). These are advantageous because the networks of simple processing units are learning systems that can acquire the structure of cognitive domains through training. Additionally, they contain computational parameters that alter the efficiency of learning, and so provide a tool to explore non-optimal conditions for acquisition. In the following paragraphs, I discuss four different theoretical issues I and various colleagues have investigated using connectionist modeling.

5.1 The contribution of the developmental process to producing behavioral impairments

In one model, we explored the implications of damaging a learning system in its initial state (analogous to a developmental disorder) compared to damaging a system in its trained state (analogous to an adult acquired deficit) as a way of gauging the potential contribution of a developmental process to generating behavioral impairments (Thomas & Karmiloff-Smith 2002a). The results demonstrated that some types of damage hurt an information processing system much more in its 'adult' state (e.g., severing network connections) while others hurt the system much more in the 'infant' state (e.g., adding noise to processing or blurring the input). The adult system can tolerate noise because it already has an accurate representation of the knowledge, but loss of network structure leads to a decrement in performance since connections contain established knowledge. By contrast, the infant system can tolerate loss of connections because it can organize remaining resources to acquire the knowledge, but the infant system is impaired by noisy processing because this blurs the knowledge that has to be acquired. This result echoes the conclusion of McDonald (1997) that a key factor in predicting the success of language acquisition across typical and atypical populations is whether the child has good representations of speech sounds.

5.2 Case study: English past tense formation in Williams syndrome

In other work, we have applied connectionist models to a much more detailed, data-driven consideration of one domain and one developmental disorder, the acquisition of English past tense formation in Williams syndrome (Thomas

& Karmiloff-Smith 2003). The model combines lexical-semantic information about a verb with phonological information about the verb's stem to generate its past tense form (Thomas & Karmiloff-Smith 2003; see Lavric, Pizzagalli, Forstmeier, & Rippon 2001 for discussion of this architecture). It thus allows detailed consideration of the relative influence of lexical-semantic and phonological constraints on the acquisition of this aspect of morphosyntax. As an outcome of the *normal* developmental process, the network comes to rely differentially on the two sources of information for driving two types of inflection, regular past tenses (talk ⇒ talked, wug ⇒ wugged) and irregular past tenses (go ⇒ went, hit ⇒ hit, think ⇒ thought). In particular, the system relies more heavily on lexical-semantic information for driving irregular inflections, so that in the trained model, a lesion to lexical-semantics differentially impaired irregulars (see also Joanisse & Seidenberg 1999). Our simulations focused on a cross-sectional developmental trajectory for the acquisition of regular, irregular, and novel verb past tense formation that we had generated from around 20 individuals with WS and 50 control children and adults (Thomas et al. 2001). These data indicated that individuals with WS exhibited a delay in the acquisition of the English past tense that was equal for regular and irregular verbs, but also a reduced tendency to generalize known inflectional patterns to novel verb forms.

We then set out to explore whether alterations to the model's initial constraints could account for these three features of the WS data. As we have seen, various claims have been made that there are subtle differences in the language system of individuals with Williams syndrome, including the proposals that their phonological representations may be atypical and perhaps rely on sensitive auditory processing, that their semantic representations may be atypical, or that semantic information about words may integrate poorly with phonology. Having established that the model could capture the normal developmental trajectory in this domain, we altered the initial constraints of the untrained network model to implement each type of proposed deficit. The results revealed that a manipulation of the phonological representations that reduced their similarity and redundancy was sufficient to reproduce the delay for regular and irregular past tense forms, as well as the reduction in generalization. Second, the pattern could also be produced when noise was added to the information coming from the semantic system during the acquisition of the past tense. Third, elimination or weakening of the semantic contribution produced a pattern inconsistent with this set of WS data comprising a selective delay for irregular verbs and no reduction in generalization (though see Clahsen & Almazan 1998, for a report of this pattern in a small sample of

4 children with WS). Lastly, slowed learning failed to produce a reduction in generalization, suggesting that delayed development alone was insufficient to explain WS performance and that atypical computational constraints are likely to be involved. This detailed modeling work was therefore able to test the viability of several competing hypotheses on the causes of particular language impairments in Williams syndrome. Manipulations to phonology or to the integration of phonology and semantics were able to simulate the past tense data; manipulations to semantics alone or delayed development were not.

5.3 Domain-specific versus domain-general deficits: A possible approach to explaining behavioral impairments in SLI

In a wider exploration of the model described above, we found that altering a 'domain-general' internal computational constraint prior to exposure to the problem domain could change the network's balance between the way it exploited lexical-semantic and phonological information during learning (Thomas in press b). With this atypical parameter setting, the network generated a profile of performance on English past tense acquisition that is not dissimilar to that reported for children with SLI. For example, van der Lely and Ullman (2001) reported that in a past tense elicitation task, children with SLI showed low levels of inflection for both regular and irregular verbs (10–20% correct) and similarly low levels of extension of the regular rule to novel stems. Since regulars are normally inflected more accurately than irregulars, this amounts to a greater deficit for regular verbs – one might view this as a kind of developmental fractionation. Van der Lely and Ullman's explanation of this pattern of behavior relies on a linguistic theory that distinguishes separate mechanisms for acquiring regular and irregular verbs (Pinker 1991). Regulars are learned by a rule-implementing mechanism whereas irregulars are learned by an associative memory (see Ullman & Pierpont in press, for a similar account where the two mechanisms are aligned with procedural and declarative memory systems in the brain). According to Ullman and colleagues, the children with SLI are unable to learn the regular rule due to an initial impairment in their rule-based/procedural system and the few regulars and irregulars that are correctly inflected reflect the compensatory action of the associative/declarative system. The idea that regulars are now inflected by a compensating associative memory system instead of a rule mechanism in the SLI group is supported by evidence of abnormally large frequency effects for regular verbs – frequency effects are taken to be the hallmark of domain-general associative memory.

It is important to be clear about the chain of inference in this case, because it clearly illustrates how researchers can move from behavioral evidence to deducing structural fractionations of the language system. The relatively greater impairment of regular inflections, along with the increased frequency effects in residual regular inflection are taken as evidence that in SLI, there has been a startstate deficit to a *domain-specific* computational structure responsible for learning regular past tense forms. It is important because the connectionist past tense model was able to simulate the same behavioral data without postulating any domain-specific fractionation, and moreover, exhibit the behavioral pattern as the product of an implemented developmental process.

To understand how the model simulated these data, we need to understand a little more about it. The model employs a 'three-layer' architecture, where a layer of internal processing units intercedes between the input layer (in this case representing lexical-semantics and verb-stem phonology) and the output layer (here representing inflected verb phonology). This internal or hidden layer is a common representational resource involved in processing regular, irregular, and novel inflections. The manipulation we applied to the network was to alter the initial properties of these hidden layer units. In particular, we reduced the sharpness of their thresholding functions. This manipulation roughly had the effect of attenuating the 'discriminability' of the units, making all computations fuzzier. The network was less able to learn sharp category boundaries in the problem domain to which it was exposed, requiring far more training than normal to generate these discriminations.

When the disordered network was 'aged-matched' to a normally developing past tense network, it exhibited low levels of regular and irregular inflection, along with poor regularization of novel stems. In other words, the disordered network gave an approximate fit to the SLI data presented by van der Lely and Ullman (2001). Importantly, in the model just as in the empirical data, regular verbs now exhibited an elevated frequency effect. Subsequent analysis of the network revealed that this was because regular inflection was being driven more strongly by lexical-semantic input than in the normal network. In effect, the system was treating regulars in the same way as irregulars, as if all verbs were exceptions to be generated via support from the lexicon.

On the face of it, this model would appear to parallel van der Lely and Ullman's explanation of their SLI data: residual regular inflection reflects the action of the declarative memory system storing word-specific information. Similarly, regulars and irregulars were treated in the same way in the disordered network, with equivalent reliance of lexical-semantics and equivalent sized frequency effects. Crucially, however, the startstate manipulation to the

connectionist network was not to a domain-specific processing structure affecting only regulars, as assumed by Ullman and Pierpont and van der Lely. Instead, the computational manipulation targeted a general processing resource used to inflect both regular and irregular verbs. However, the particular computational property altered was one upon which regular verbs differentially relied, since such verbs differ yet must all be treated in the same way. This requires sharp category boundaries that delineate regular space in which all items will be treated the same. The alteration of this domain-relevant property was a deflection of the developmental trajectory such that in terms of the relative size of deficits, there was an apparent fractionation between regular and irregular verbs. These initial alterations to the common computational resource had the effect of altering the balance of the information sources on which the network relied to generate past tense forms. Phonological regularities were downplayed, while word-specific information was emphasized. The atypical constraints of the learning system served to alter the interaction between phonological and semantic sources of knowledge during development of this morpho-syntactic ability.

In sum, this modeling results demonstrates that behavioral evidence taken by van der Lely and Ullman (2001) and Ullman and Pierpont (in press) to indicate a structural fractionation of the language system in SLI could also be explained in terms of a learning system without such a fractionation, and the initial manipulation of a computational parameter with no specific reference to regular or irregular verbs.

5.4 Inferences from the comparison of developmental profiles from different disorders

Modeling work also sheds light on the interpretation of similarities and differences in the way different disorder groups acquire language. The fact that we can take a model of normal development and create developmentally impaired systems of various types (noisy systems, systems with memory impairments, slow learners, and so on) allows us to explore the extent to which qualitatively different behavioral profiles are generated by altered internal constraints. We explored this in two recent models: the past tense model already discussed and a model of syntax acquisition.

One explanation of the similarities identified between the developmental profiles and patterns of errors across different disorders is that these similarities reflect immovable internal constraints of the language learning system (Newport 1990). The notion of a 'developmental delay' is predicated on iden-

tifying such similarities in children that are not reaching the landmarks at the correct ages. It is deployed even in the case where mastery is never reached. Similarities may therefore be taken to imply that nothing is qualitatively different in the system: it is just not 'working very well'. However, it is also possible that similarities between typical and atypical development have another explanation: the range of behaviors that individuals can exhibit in language development is constrained by the common physical, social, and informational environment in which each individual's cognitive system is embedded. More specifically, behaviors normal or otherwise are in part constrained by the structure of the problem domain to which the cognitive system is exposed, whatever its underlying architecture. The extent to which cognitive architecture is *visible* in the behavioral changes and error patterns exhibited across development is a serious and unresolved issue. The simplest illustration of this idea is a cognitive domain that has an easy part and a hard part. A wide range of learning systems would naturally acquire the easy part before the hard part. Consequently, a common developmental profile here would tell us little about the actual learning system involved. To investigate this proposal, we exposed a variety of associative architectures to the past tense domain, varying the computational resources that the learning system brought to the problem (Mareschal et al. forthcoming). The results indicated that there was indeed great variation across the developmental profiles. However, the systems also exhibited similarities in their profiles. In particular, regular verb acquisition was usually in advance of irregular acquisition, and generalization of the regular rule was usually weaker to novel stems that rhymed with irregulars than to those that did not. These patterns were a result of the structure of common past tense domain that each model learned, including the similarities between verbs and type and token frequencies of the various items involved.

Dick et al. (2001, 2004) recently argued that similarities in syntactic deficits found in adults with aphasia and in children with developmental language impairments can also be traced to features of the shared problem domain. In particular, in a comprehension task (agent-patient role assignment), low frequency constructions and non-canonical subject-object word order constructions such as passives and object clefts ('the cat was chased by the dog', 'it was the cat that the dog chased') revealed greater behavioral impairments than high frequency and canonical order constructions like actives and subject clefts ('the dog chased the cat', 'it was the dog that chased the cat'). We trained a recurrent, sequence processing connectionist network on sentences of this form in the frequency that young children hear them (Thomas & Redington 2004). The network had to perform the same comprehension task as human

subjects, identifying the agent in each sentence. The trained network showed the normal adult pattern of difficulty across the constructions. When it was trained with an initially reduced level of computational resources, it was also successful in simulating the exaggerated pattern of difficulty shown by children with developmental disorders. Importantly, the model also demonstrated relatively less vulnerability of constructions learned on the basis of unique lexical cues (such as passives, indicated by the word 'by') and relatively more vulnerability of constructions learned on the basis of sequence cues (such as object clefts, indicated by the two nouns that are not split by an intervening verb). The behavioral data (Dick et al. 2004) were also consistent with this differential effect. This pattern emerges in the model because reducing the initial computational resources produces a greater impairment in analyzing global information across sentences than in analyzing local information from individual lexical items. The consequence is that although the structure of the task domain paints a broad picture of task difficulty, the strengths and weaknesses of the computational learning system modulate this pattern.

In sum, models of two different aspects of grammar acquisition demonstrate that some similarities between atypical and normal development are the consequence of the problem domain. Disordered learning systems only serve to modify this pattern, sometimes in subtle ways. This line of computational work indicates firstly that the attributions of language disorders to 'developmental delay' on the basis of an absence of 'qualitative' differences need to be treated with caution; and secondly, the inference that behavioral similarities across different populations reflect internal constraints is not a secure one – they may as easily reflect external constraints.

6. Conclusions

In this chapter, we have considered how atypical profiles of language impairments may be informative about language acquisition. I have argued that the appropriate framework for explanations of deficits in developmental disorders is in terms of constraints on the developmental process – whether a given theory assumes the presence of domain-specific modular structure prior to language acquisition or assumes that such structure is the product of the developmental process itself. We considered characteristics that the atypical developmental process should incorporate such as interactivity and compensation. If these characteristics do not figure within the developmental process, theories must explicitly stipulate why they should not occur. The example of Williams

syndrome was used to illustrate how researchers can begin to identify the particular constraints that have deflected language development in an atypical population. Finally, computational modeling of atypical language acquisition was discussed, both as a method for testing whether a given set of atypical constraints (such as the balance of phonological and lexical-semantic information) are sufficient to generate particular behavioral impairments, and also as a way to assess the strength of inferences drawn from behavioral data. In the case of the latter, we saw how modeling indicated that behavioral dissociations in language development do not necessarily imply underlying structural fractionations, and how behavioral similarities between typical and atypical language acquisition do not necessarily stem from shared internal constraints but from the structure of the problem domain. Finally, we must note the context of this research. Understanding the constraints that shape and deflect the acquisition of language is an important step towards understanding how we may intervene to optimize the outcome of language learning in atypical populations.