

The biological basis of language: insight from developmental grammatical impairments

Heather K.J. van der Lely and Steven Pinker

Department of Psychology, Harvard University, William James Hall 970, 33 Kirkland Street, Cambridge, MA 02138, USA

Specific language impairment (SLI), a genetic developmental disorder, offers insights into the neurobiological and computational organization of language. A subtype, Grammatical-SLI (G-SLI), involves greater impairments in ‘extended’ grammatical representations, which are nonlocal, hierarchical, abstract, and composed, than in ‘basic’ ones, which are local, linear, semantic, and holistic. This distinction is seen in syntax, morphology, and phonology, and may be tied to abnormalities in the left hemisphere and basal ganglia, consistent with new models of the neurobiology of language which distinguish dorsal and ventral processing streams. Delineating neurolinguistic phenotypes promises a better understanding of the effects of genes on the brain circuitry underlying normal and impaired language abilities.

Developmental disorders as a window into the biology of language

Given the lack of animal models for language, and the inability to use invasive procedures with humans except out of medical necessity, our knowledge of the neurobiology of language has long depended upon natural experiments. During the 19th and 20th centuries, studies of patients with acquired brain lesions provided key insights [1–3]. Understanding of language in the 21st century promises to be enriched by data from developmental disorders. SLI, a family of language impairments in otherwise normal children, is highly heritable and has been linked so far to four genes. These discoveries provide a new route to understanding the complex pathways from genes and environment to the neural systems underpinning language.

This understanding depends, however, on breaking down the coarse categories of ‘language’ and ‘language impairment’ and examining the way that specific components of language are affected in specific disorders, and how they correlate with brain function and structure. That is, rather than searching for a direct link from genotype to behavior, we suggest linking genetic variants

with alterations in the neural substrates of subcomponents of language processing.

Specific language impairment (SLI)

SLI is a heterogeneous family of impairments which affect the acquisition of language in 7% of children, an average of two in every classroom [4]. It frequently co-occurs with other disorders such as dyslexia, autistic spectrum disorders, and attention deficit and hyperactivity disorder [5,6], with which it also shares some phenotypic and genotypic characteristics [7,8]. Many genetic variants contribute to SLI across individuals, consistent with the heterogeneity of the disorder [7]. Despite this heterogeneity, the majority of children are impaired in grammatical functions, particularly syntax and morphology, and often phonology as well (Box 1) [4,9,10]. One of us has identified a subtype of the broad SLI category called G-SLI, which is concentrated in grammar, though it may embrace secondary deficits, for example, in the lexicon ([4,9–11]; but see also [12–14]). Though it was discovered in English-speaking children, G-SLI has since been identified in other languages [15–17].

Grammatical phenotypes of SLI

This review focuses on what G-SLI can reveal about the structure and neural instantiation of language. Crucially, G-SLI is not a global impairment of language or even of grammar, but is strongly manifested in certain aspects of linguistic performance while leaving others largely intact. This raises the possibility that the contrast reflects a key division within the neural or genetic substrates of language. In particular, children with G-SLI have difficulty interpreting and producing syntactic structures such as ‘wh’-questions, the passive voice, and tense-marking; words that must be grammatically inflected in real time; and complex phonological structures embracing multiple syllables and clusters of phonemes. However, they have age-typical performance in syntactic tasks in which lexical semantic information is sufficient; in morphological tasks in which stored, nondecomposed forms are sufficient; and in phonological tasks in which strings of phonological units are sufficient. The problems with composed forms, moreover, persist into adulthood. We suggest that this pattern of deficits may reflect two modes of grammatical representation and processing which we call Basic and Extended (Box 1), and that individuals with G-SLI are specifically impaired in processing Extended representations.

Corresponding author: Pinker, S. (pinker@wjh.harvard.edu).

Keywords: specific language impairment; neurolinguistics; genetics of language; brain and language; language disorders.

1364-6613/

© 2014 Elsevier Ltd. All rights reserved. <http://dx.doi.org/10.1016/j.tics.2014.07.001>

Box 1. Components of language

Language is traditionally divided into several subsystems [79]:

- **Syntax:** The combination of words into phrases and sentences, and assignment of grammatical relations (subject, object, head, etc.) which determine their compositional meaning.
- **Morphology:** The combination of words or parts of words (morphemes) into new words, further subdivided into 'inflection' (modifying a word according to its role in the sentence) and 'derivation' (creating a new word from old ones).
- **Phonology:** The combination of sounds into morphemes, and the modification of sounds according to their contexts.
- **Pragmatics:** Principles governing the use of language in a discourse and communicative context.
- **Lexicon:** The component of memory which stores words, idioms, and other fixed forms.

We concentrate on the first three, and propose that they are cross-classified by a distinction in representation and processing:

Extended grammatical representations are:

- abstract, consisting of categories defined by their grammatical privileges rather than their semantic content;
- hierarchical, defined by a tree of constituents embedded in larger constituents;
- nonlocal, potentially spanning long distances in the string;
- composed, namely assembled into meaningful combinations by rules.

Basic grammatical representations are:

- semantic and lexical, consisting of words or features of meaning;
- linear, defined by left-to-right ordering;
- local, involving adjacent or nearby elements;
- holistic, consisting of entire assemblies stored in memory.

Extended versus Basic syntax

Extended syntax involves hierarchical structures and dependencies between words, often spanning the entire clause, which are computed in real time. For example, in 'wh'-questions such as 'Who did Joe see __?', the 'wh'-word and the empty position after 'see' are in a dependent relation, which may be analyzed as the movement of the word from its original position in an underlying structure (Figure 1). Additionally, Extended syntax is abstract: the assembly and interpretation of phrases depend on their grammatical categories (noun, verb, tense) and relations (subject, head, complement), each defined by a pattern of intercorrelated privileges (where they may occur, how they may be inflected, what can substitute for what). Abstraction is central to grammar: the acquisition of abstract symbolic rules enables a person to generalize a pattern learned from a finite number of exemplars to an infinite number of new ones which need not resemble them in sound or meaning [18,19].

Basic syntax, in contrast, involves relations between words that can be determined from the meanings of the words themselves or from dependencies between adjacent words. Basic syntax may consist of holistic representations, in which sequences are stored and retrieved without necessarily analyzing their grammatical structure, and instead are linked directly to their semantic and pragmatic properties and their ordering relative to adjacent units.

Extended and Basic representations differ in their processing requirements. In Basic syntax, words and their features (number, gender, meaning) can be inserted directly from the lexicon, whereas in Extended syntax, relations between words within and across hierarchical units must be computed by operations such as movement and feature checking or unification. Figure 1 shows some

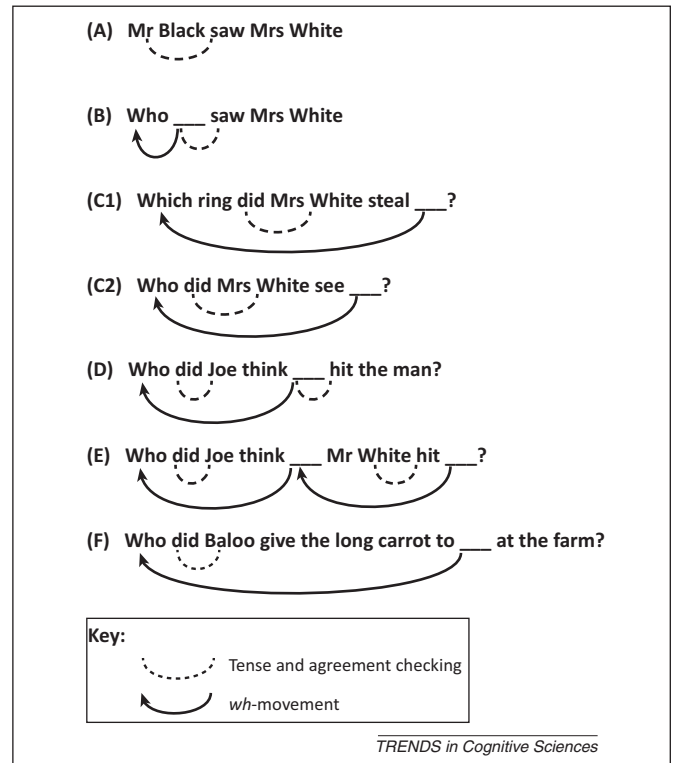


Figure 1. Examples of Extended syntax. The English rules for marking tense and agreement (A–F) and forming questions are complex and belong to what we call Extended syntax. Questions with 'wh'-words (who, what, which, etc.) such as those in (B–F) require movement (unbroken arrows) from an underlying position (underscore, cf. (A)) to a position in the surface string. The result of the movement is audible when the 'wh'-word corresponds to an object (C), the subject of an embedded clause (D,E), or the object of a preposition (F), but it is covert when the 'wh'-word corresponds to the subject of a main clause (B). Tense and subject-verb agreement are obligatory in English main clauses; in 'wh'-questions they must be marked on the auxiliary 'do' if the clause lacks any other auxiliary, and the auxiliary must be inverted with the subject (Did he leave? rather than Left he?). These operations are handled by a complex sequence of movement and feature-checking operations, which we abbreviate here with broken lines. If checking or movement has not been reliably computed, as (we hypothesize) is common in G-SLI, then for verbs that should be marked for tense, the infinitival form may be used instead, and the auxiliary may be omitted. For 'wh'-questions, the problem may be manifested as a filled gap or, if the movement is partial, the 'wh'-word may be copied in the medial moved position (D, E). See Table 1 for examples of errors produced by such children.

of the extended syntactic relations which must be computed, according to a major theory of grammar [20], in assigning tense to a clause and in producing or interpreting 'wh'-questions; other theories require operations of comparable complexity.

Impairments in syntax

Table 1 shows the results of a variety of experiments, differing widely in their methods and processing demands, in which children with G-SLI display problems with Extended syntax (specifically, 'wh'-questions similar to those illustrated in Figure 1), but perform well in control conditions requiring only Basic syntax or lexical semantics [9,10,21].

Experiments testing other syntactic constructions show similar patterns. For example, teenagers with SLI rely on number information as a shortcut to understanding relative clauses in the same way that unimpaired six-year-olds do [22]: upon hearing 'The cat that is chasing the dogs is black', they understand that the cat is black rather than

Table 1. Patterns of errors with ‘wh’-questions by language-impaired children. Table 1 summarizes experiments with G-SLI children investigating ‘wh’-questions (who, what, which) in production, grammaticality judgments, on-line cross-modal priming, and EEG recordings. In each experiment the children failed to compute the extended syntactic relations between the ‘wh’-word and the position to which it is related (Figure 1). The diversity of methods and tasks suggest that the deficit is independent of the paradigm and processing demands. Similar patterns are found with French, Italian, German, Hebrew and Greek [15–17,80,81]

Task ^a	Description	Lead-in or prime (see Figure 1 for the syntactic structure)	Control child’s typical response	Child with G-SLI’s typical response
Elicitation of ‘wh’-questions	‘wh’-subject question	Someone saw Mrs White. Ask me who.	Who saw Mrs White?	Who someone see Mrs White?
	‘wh’-object question	Mrs White stole a ring. Ask me which one?	Which ring did Mrs White steal?	Which [...] Mrs White stole the ring?
Sentence Judgments	Semantic judgment	(i) Barbie bakes the cakes (ii) Barbie drinks the cakes	(i). Yes (ii). No	(i). Yes (ii). No
	Syntactic judgment: Tense	(i) Who did Mrs White see? (ii) Who did Mrs White saw?	(i). Yes (ii). No	(i). Yes (ii). Yes
	Syntactic judgment: ‘wh’-movement	(i) Who did Mrs White see? (ii) Who did Mrs White see someone?	(i). Yes (ii). No	(i). Yes (ii). Yes
Elicitation of Embedded ‘wh’-questions	Embedded subject ‘wh’-question	Joe thought someone hit the man. Ask me who?	Who did Joe think hit the man?	Who did Joe think who hit the man?
	Embedded object ‘wh’-question	Joe thought Mr White hit someone. Ask me who?	Who did Joe think Mr White hit?	Who [...] Joe think Mr White hit someone?
Cross-modal Sentence–Picture Priming	The child is shown a picture of a rabbit or a ladder at one of three points during a spoken sentence and presses a button indicating whether it is animate or inanimate. Response time for position [2], after the adjective, serves as a control. Priming (faster responding) at position [1] shows sensitivity to basic verb semantics (<i>give</i> + recipient). Priming at position [3] (the ‘wh’-gap) shows sensitivity to syntactic movement: the rabbit matches the ‘who’ which had been moved from that position.	Baloo gives a long carrot to the rabbit at the farm. Who did Baloo give [1] the long [2] carrots to [3] at the farm?	[1] No priming. [2] No priming. [3] Priming.	[1] Priming. [2] No priming. [3] No priming.
Event Related Potentials (ERP)	Violation of expectation based on semantics (ERP time-locked to the underlined word)	(i). Control: Barbie bakes the <u>cakes</u> in the oven. (ii). Violation: Barbie bakes the <u>people</u> in the oven.	N400	N400
	Violation of expectation based on syntax (ERP time-locked to the underlined word)	(iii). Control: Who did Barbie push the <u>ball</u> into? (iv). Violation: Who did Barbie push the <u>clown</u> into the wall?	ELAN, P600	Right lateralized N400, P600

^aThe Elicitation task [82] shows that children and teenagers with G-SLI fail to compute filler-gap relations and fail to check and mark tense (Figure 1). In the Sentence Judgment task [83], G-SLI teenagers correctly reject semantic violations and accept grammatical sentences, but incorrectly accept sentences with errors in filler-gap relations and tense. The task Eliciting Embedded Questions (A. Archonti, PhD thesis, University College London, 2003) illustrates omission of the auxiliary ‘did’ and a copying error, where the ‘wh’-word is only partially moved and ends up in an intermediate position. The Cross-Modal Priming study [25] shows that in children and teenagers with G-SLI, a ‘wh’-filler does not prime its related word at the gap position, though it does prime it at the verb’s offset, suggesting they are sensitive to the verb’s semantics. The ERP study [21] shows that G-SLI teenagers have an appropriate response (the N400) to violations of lexical semantics (i and ii). Responses of violations to syntax were tested with ‘wh’-questions in which agents perform actions on inanimate objects. Unimpaired subjects perceive a syntactic violation as soon as they encounter an animate noun that could be linked with the ‘wh’-word, such as *clown* in (iv), and thus show an early syntactic brain response (ESBR); the G-SLI subjects instead show the N400 ordinarily elicited by semantic violations.

the dog, because plural ‘dogs’ cannot be the subject of singular ‘is black’; whereas with ‘The cat that is chasing the dog is black’, which provides no such cue, their performance is close to chance. Other examples of structures which require Extended syntax for interpretation, and with which individuals with SLI show comprehension difficulties, are reversible passive sentences, such as ‘The truck was hit by the car’ [23]. Likewise, they have trouble assigning an ambiguous pronoun to an antecedent specified by grammatical constraints, as in ‘Mowgli said

Baloo is tickling him’ (in which Baloo must be tickling Mowgli, not himself, as opposed to ‘Mowgli said Baloo is tickling himself’, where Baloo can only be tickling himself), while having no trouble assigning pronouns to antecedents of the same gender, as in ‘Jane says Paul is tickling her’, where ‘her’ can only be female ‘Jane’, not male ‘Paul’ [24]. Similar patterns are found in SLI children speaking other languages [16]. When the memory load required to link a pronoun to its antecedent is controlled or reduced (as in agentless passive sentences such as ‘The truck was hit’), or

when memory abilities are matched in a control group, the individuals with G-SLI still have more trouble with sentences requiring Extended syntactic analysis [9,22,25]. G-SLI children's difficulties are also seen in their production, such as the omission of tense-marking (as in 'Yesterday I fall down' [9,13], a phenomenon at the intersection of syntax and morphology).

The abstractness of Extended syntax also poses problems. When teenagers with G-SLI try to learn simple abstract rules, such as ABB versus ABA (realized in contrasts such as 'vi ko ko' versus 'vi ka vi'), they failed to learn the instances, to generalize them to new sequences, and to show the expected electrophysiological brain response to violations.

Extended versus Basic morphology

G-SLI also impairs morphology in a selective fashion which targets abstract combinatorial operations. Extended morphological objects are generated, perceived, and represented as combinations of two or more morphemes, whereas Basic ones are stored and retrieved as wholes, or at best as containing parts which do not freely combine with other parts [9,10]. For example, in English, irregular past-tense forms ('ate', 'bought') are basic and handled by memory, whereas regular past-tense forms ('jumped', 'agreed') are morphologically complex and may be computed from the verb stem and '-ed' suffix, particularly when the verb is rare or novel [19,26].

Impairments in morphology

Problems with regular inflectional morphology (an Extended process) are a hallmark of SLI in English speakers [9,27–30]. In contrast to unimpaired children, children with G-SLI seem to habitually store regular forms as wholes. They have more trouble with low-frequency than with high-frequency regular forms (suggesting they have memorized the high-frequency ones). They are no better at producing regular than irregular forms. They produce regular plurals inside compounds ('rats-eater'), which unimpaired speakers deem ungrammatical. Finally, they have great difficulty inflecting novel verbs and nouns, where the demand on computation is highest and memory lookup is unavailable [9,10,28–30].

Even when decomposition of extended morphological forms would facilitate production and perception, children with G-SLI link the word's sound directly to its lexical form in memory [9,10]. For example, words ending in sequences that never occur in monomorphemic forms, such as 'vd' in 'loved' and 'gd' in 'hugged', or the German 'fst' as in 'du kaufst' (you buy), must have a morphological boundary falling within the word-final cluster, making them unambiguously inflected. Unimpaired individuals are sensitive to this cue; G-SLI individuals are not [9,10,31].

At first glance, derivational morphology (Box 1) ought to parallel inflectional morphology, with predictable, freely generated regular forms (e.g., 'uncorkability') and idiosyncratic, stored irregular ones (e.g., 'destroy–destruction'). Nonetheless, with unimpaired speakers, even regular derived forms seem to be stored and retrieved as wholes [32], presumably because derivational suffixes have more semantic content than inflectional ones and their combinations are

more likely to be lexically specific. We speculate that this may parallel the surprising ability of G-SLI individuals to provide regular derived forms on demand, such as 'sillier' from 'silly' and 'rocky' from 'rock' [33]. Although the children rarely omit derivational suffixes, they occasionally misapply them, as when they convert 'happy' to 'happer' or 'happes' rather than 'happier' and 'happiest', and 'frills' to 'frillys' rather than 'frilly'. This pattern may reflect their mosaic of abilities: they have slots in memory for familiar derived forms, but are impaired at reconstructing the sequence of morphemes that fills it.

Extended versus Basic phonology

We now heuristically apply the Extended–Basic distinction to a third component of language, phonology. In this extension, Basic phonological representations consist of strings of phonemic segments made up of distinctive features. Extended phonological representations are hierarchically organized into prosodic units of increasing size: the syllable, the foot, the prosodic word, the intonational phrase, and the utterance [34].

In speech perception, the Basic and Extended aspects of phonological representations are initially processed via independent streams which are then later integrated [35,36]. Basic representations require the analysis of rapidly changing events in the speech signal within a narrow time window. This enables acoustic distinctions to be drawn between phonemes (such as the place contrast between 'b' and 'd') and phoneme sequences (such as 'ts' versus 'st'). Basic phonology is necessary for building segments into syllables.

In contrast, Extended representations require the analysis of more slowly unfolding events in the speech signal using a wider sampling window. Extended phonology is necessary for building syllables into feet, words, and larger prosodic domains. Amongst other things, it handles word stress, syllable rate, and intonation.

Impairments in phonology

Contrary to a common assumption that SLI originates in a difficulty in processing transient speech signals, which then propagates to higher-level representations, individuals with G-SLI have few problems processing Basic phonological representations such as segments. They are age-appropriate in categorical perception of phonemes such as 'b' and 'd' [4,37], and children with broad SLI can distinguish minimal segmental contrasts between stored words, such as 'van' and 'fan' [37]. Children with broad SLI also compensate for coarticulation effects that can mask phonemic contrasts; for example, they recognize that 'brow[n] roof' and 'brow[m] bell' contain instances of the word 'brown' but that 'brow[m] lamp' does not [38]. They can also follow instructions to exchange segments in words, such as converting 'sad cat' to 'cad sat' [37].

Basic phonology is also sufficient for people with broad SLI to perceive and articulate adjacent sounds in consonant clusters in real and novel words. Their problem consists in putting them in the wrong structural position, saying 'flakesta' rather than 'faklesta', presumably because they fail to represent details of the extended phonological structure [39]. Similarly, they have more trouble producing

regularly inflected forms that end in complex clusters ('jumped') than those with simple endings ('sewed') [40]. This may explain their reduction of trisyllabic derived forms such as 'happiest' to 'happesst' [33].

Individuals with G-SLI also have difficulties with the Extended representations needed to handle prosody. In non-word repetition tasks, they make more errors in producing consonant clusters as the words become more prosodically complex [41]. Children with broad SLI also have difficulty using prosodic cues to distinguish phrases such as '[red] [and black and pink socks]' from '[red and black] [and pink socks]' [42].

Neural substrates of the distinction between Basic and Extended grammar

Recent advances using electroencephalography (EEG), magnetoencephalography (MEG), functional MRI (fMRI), and diffusion tensor imaging (DTI) tractography have led to sophisticated new models of the neural organization of language [35,43–45]. They allow us to assess whether the Extended processes, which pose the greatest problems in G-SLI, implicate brain networks that are different from those supporting Basic processes.

Syntax

The new models (Figure 2) go beyond the classic Broca's and Wernicke's areas and distinguish at least three fronto-temporal networks related to syntactic processing [43]. First, a dorsal pathway links Brodmann area 44 (a part of Broca's area) via the arcuate fasciculus (AF) to the posterior superior temporal gyrus (STG; a part of Wernicke's area). It is said to underlie 'hierarchical structures independent of semantics' in natural and artificial languages, and 'syntactically complex sentences,' particularly those involving movement [43,44]. It is associated with the rapid (100–200ms) early left anterior negativity [44] and

syntactic-mismatch negativity [46–48], which we will lump together as early syntactic brain responses (ESPRs).

There are also two ventral pathways. One, which supports 'local phrase structure,' links the frontal operculum (adjacent to Broca's area) via the uncinata fasciculus (UF) to the anterior STG, and it is associated with a later (300–700ms) left anterior negativity (LAN). The other, supporting retrieval of stored words and lexical semantic processing, links BA45 (another part of Broca's area) via the extreme capsule fiber system (ECFS), to the posterior temporal lobe, particularly the STG and superior temporal sulcus (STS).

The functions of the dorsal and ventral syntactic circuits correspond well to what we call Extended and Basic syntax. Interestingly, the dorsal pathways, unlike the ventral ones, do not fully mature until approximately 7 years of age [49]. Accordingly, the brain responses of 2-year-olds are sensitive to sentences that violate Basic syntax, with a left lateralized ERP, but not to violations of Extended syntax [50], which only develops by 7–9 years of age [21]. Moreover, the dorsal pathways in human brains differ substantially from those in other primates, suggesting that phylogenetic changes to the dorsal pathway may have been a key driver of the evolution of language [43].

A still later neural response, which is elicited with both extended and basic syntax, reaches a maximum at approximately 600ms (the P600), and is observed bilaterally in frontal or parietal–central locations. It represents a second-pass reanalysis under participants' strategic control [51]. The P600 also involves subcortical structures, namely the caudate nucleus of the basal ganglia, which responds to ungrammatical and ambiguous sentences [52].

Morphology

Two aspects of morphology require Extended processing: the recognition that a word must be inflected in certain

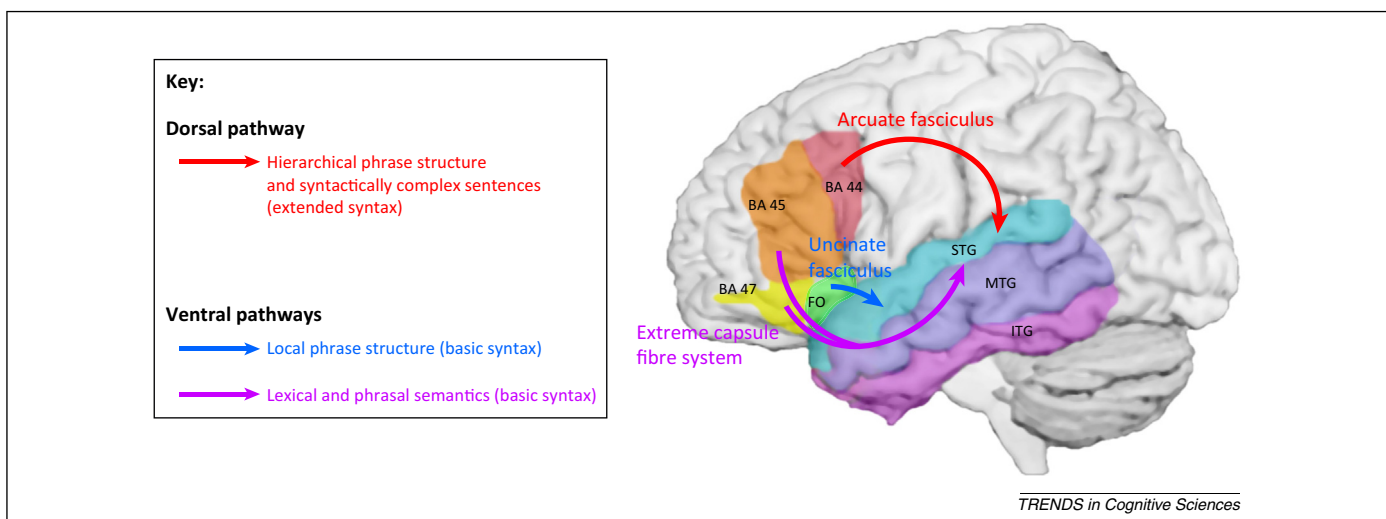


Figure 2. Neural correlates of Extended and Basic syntax. Syntactic processing in the brain is implemented in distinct dorsal and ventral circuits which may correspond to Extended and Basic syntax [43,44]. The dorsal route (unbroken red arrow) links Brodmann Area 44 (BA44, a part of Broca's area) via the arcuate fasciculus to the posterior superior temporal gyrus (STG, a part of Wernicke's area); this pathway has been implicated in complex syntactic processing, including hierarchical phrase structure and movement, that is, Extended syntax. The caudate nucleus of the basal ganglia (not shown), a subcortical structure, is interconnected with frontal cortex, and it has also been found to affect Extended syntax. The first of the two ventral circuits (blue arrow) links the frontal operculum (FO, the cortex inferior and medial to BA 44 and 45, mostly hidden) via the uncinata fasciculus to the anterior STG; it supports local phrase structure. The second (purple arrow) links Brodmann Areas 45 (BA45, another part of Broca's Area) and 47 via the extreme capsule fiber system to the middle portion of the superior and middle temporal lobe; it supports retrieval of stored words and associated semantic processing. The two ventral pathways, therefore, may correspond to Basic syntactic processing. Abbreviations: MTG, middle temporal gyrus; ITG, inferior temporal gyrus.

contexts (e.g., in tensed clauses; [Figure 1](#)); and the computation of the appropriate inflectional form, which for regular verbs (particularly novel and rare ones) involves the analysis of the word as a stem plus a suffix. Both computations are composed, hierarchical, and abstract (though local), and as mentioned, both present difficulties to children with G-SLI. Both components of morphological processing have long been associated with Broca's area [[26,53–56](#)]. A recent study using high-precision intracranial recordings revealed distinct areas within Broca's area corresponding to the selection of the appropriate tense or number (at ~320ms) and to the computation of the appropriate morphological form (at 450ms) [[57](#)]. fMRI of the same patients implicated the posterior superior temporal cortex as well. These studies suggest that, as with Extended syntax, Extended morphology is mediated by a dorsal pathway.

These studies did not distinguish regular from irregular inflection, but a comprehensive study of priming patterns in aphasic patients, together with fMRI in normal adults, specifically tied regular inflection to 'a core decompositional network linking left inferior frontal cortex with superior and middle temporal cortex, connected via the arcuate fasciculus' [[45](#)] – namely, a dorsal pathway similar to the one identified for Extended syntax ([Figure 3](#)). A recent meta-analysis of neuroimaging studies similarly found that regular inflection consistently activates BA 44 [[58](#)]. As with syntax, regular inflection also appears to engage the head of the caudate nucleus [[52–54](#)].

By contrast, the computation of irregular forms involves storage and retrieval from lexical memory, a Basic process. Studies with aphasic patients and with fMRI suggest that its neural substrate is largely distinct from that for

regulars, involving bilateral posterior temporoparietal regions [[45,53,54,59,60](#)]. They are thought to be part of a larger, complex network for lexical storage and retrieval.

Yet another pattern was found with derivational morphology. As with inflection, some derived words are regular and transparent in their structure and can be analyzed by rules ('bravely'), whereas others are opaque and irregular ('archer', 'depth'). However, regular derived forms do not seem to implicate the dorsal network, nor do they differ in their anatomical substrate from irregular ones: both are stored in a network of whole forms that activates the ventral pathway bilaterally, specifically, BA47, the temporal pole, and anterior MTG [[32](#)]. (The transparent derived forms, however, showed different patterns of activity within this network from the opaque ones.) We speculate that the nondecompositional (Basic) nature of derivation may explain why G-SLI children are competent at generating derived forms (albeit not always with the proper choice of suffix).

Phonology

A recent model of the cortical organization of phonology [[35](#)] makes a distinction similar to what we call Basic and Extended phonological processing ([Figure 4](#)). The auditory input is first subjected to spectrotemporal analyses in auditory cortex bilaterally (though with some hemispheric differences), and then to phonological analysis in the STS, before it splits into two streams. One is a left-hemisphere dorsal pathway which relates words' sounds to their articulation. This pathway is deployed in word learning, speech monitoring, verbal working memory, and the articulation of long, complex, low-frequency, and novel words. The other is a bilateral ventral pathway which relates words' sounds to their meanings, and from there

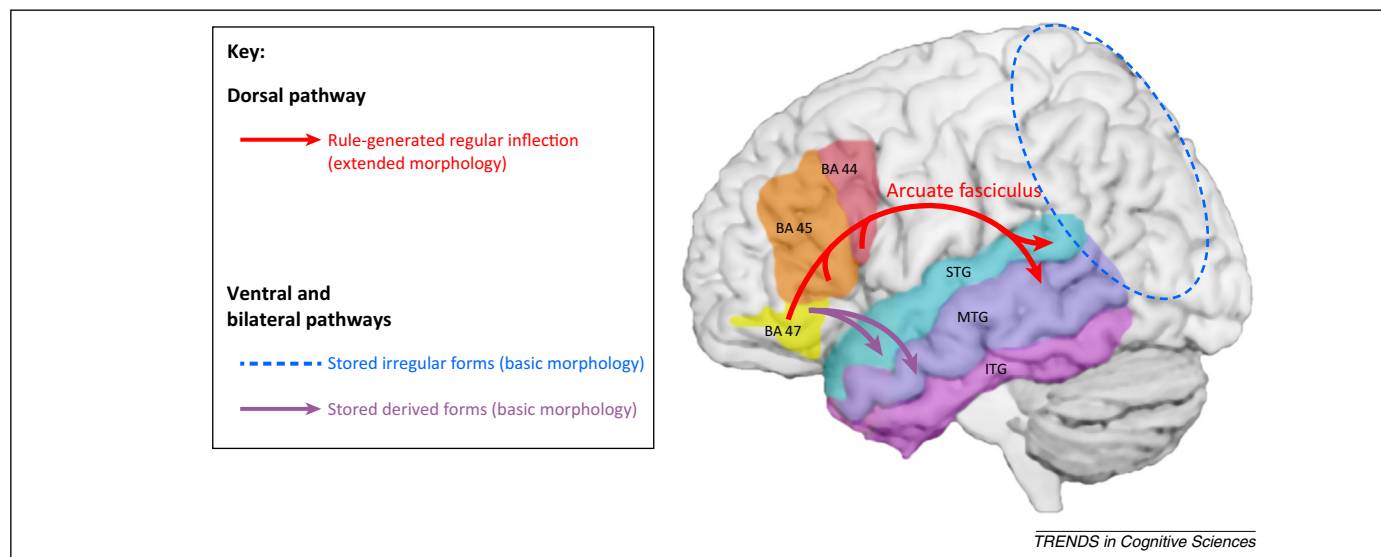


Figure 3. Neural correlates of Extended and Basic morphology. Regular inflectional forms (*walked, played*) are computed by Extended processes that closely overlap with those underlying Extended syntax, namely BA 45 extending into BA 44 and BA47, the arcuate fasciculus, and the superior and middle temporal cortex [[45](#)]. The frontal regions are part of a circuit that also includes the caudate nucleus (not shown). In contrast, the storage and retrieval of irregular forms ('brought', 'went') appears to be mediated bilaterally (blue outline) in a more diffuse set of posterior and middle temporal lobe structures [[58](#)]. Derived morphological forms, both regular ('bravely') and irregular ('archer'), activate a third, bilateral network, which we tentatively associate with the ventral pathway, specifically, BA47 extending into BA45, and the anterior superior temporal gyrus (STG) and middle temporal gyrus (MTG) (purple lines) [[32](#)]. This network may support a network of related but whole word forms. Individuals with G-SLI are impaired in productive regular inflection, an Extended process that engages the dorsal route, but their performance is less impaired with the retrieval of irregular and derived forms, a Basic process which is more tied to lexical memory, and which engages bilateral ventral and posterior routes.

Genetic correlates of grammar impairments

Language impairment is highly heritable [69,70], but as with all heritable variation, the individual genes underlying the variation have been elusive [71]. In this regard, SLI has seen more success than other psychological traits: to date, four candidate genes have been robustly associated with it.

The *FOXP2* gene on chromosome 7q, which codes for a transcription factor that regulates many other genes, is perfectly associated within a large extended family (and in a small number of unrelated individuals) with a form of SLI which co-occurs with verbal dyspraxia (a motor-speech disorder). The sequence of the *FOXP2* gene in humans differs from that in nonhuman primates, and has been identified as a target of natural selection [72].

The *CNTNAP2* gene on chromosome 7q (downstream from *FOXP2*) is associated with genetically complex forms of SLI, where different genetic variants are risk factors for the disorder [73]. The gene is also associated with autism, dyslexia and other developmental disorders that impact on language [73]. Its main linguistic manifestation demonstrated so far is in non-word repetition, which requires phonological analysis and memory, although it is also associated with low composite scores for expressive and receptive language [8,73]. Unimpaired individuals carrying the *CNTNAP2* alleles associated with SLI and autism have increased contralateral fMRI activation in the right frontal operculum (BA44) and middle temporal gyrus (BA 21), suggesting that *CNTNAP2* variants can affect brain areas associated with language, even in the general population [74].

Two additional genes that appear specific to SLI are on chromosome 16q: *ATP2C2* and *CMIP*. Both appear to cause an impairment of phonology and phonological memory [75]. Other studies have implicated other genes, with varying degrees of replicability and specificity [76]. All the genes associated with SLI appear to affect early embryonic development [7,8]. During cortical neurogenesis they are expressed in similar brain regions, including deep layers of the cerebral cortex, striatum and cerebellum, presumably leading to abnormal neurobiological pathways later in life.

A limitation of the extant literature is that most studies have used non-specific tests of language that do not isolate aspects of processing or their neural correlates. We suggest that a more promising strategy for discovering links to genes is to identify intermediate phenotypes consisting of functional and anatomical brain circuits which underlie particular aspects of the impairment (such as the distinction between Basic and Extended processing, which we suggest characterizes the G-SLI subtype). The use of quick, yet reliable tests (such as the Grammar And Phonology Screening test [77,78], which can identify G-SLI by testing extended syntax, morphology and phonology) promises to add new precision to the study of the genetic correlates of language impairment.

Concluding remarks

We have suggested that the study of language impairment has opened a new frontier of research which promises a comprehensive biological understanding of language, from

evolution and genes, through neuroanatomy and neural function, to linguistic computation, and thence to overt speech and comprehension. These advances depend on a multi-level approach. Rather than mapping genetic variants directly to overall language impairment, researchers must characterize the intermediate links by probing for finer-grained linguistic components and modes of processing, together with their neural substrates.

As a preliminary example, we have argued for a distinction between a family of Basic representations and processes and a family of Extended ones. Extended syntax involves hierarchical, composed phrase structures, abstract rules, and movement, whereas Basic syntax involves local phrases, lexical retrieval, and interpretation based on lexical semantics. Extended morphology involves multi-morphemic, composed, regular forms, whereas Basic morphology involves the retrieval of whole words comprising irregular, derived, or high-frequency regular forms. Extended phonology involves the composition of complex syllables and higher prosodic units out of segments, whereas Basic phonology involves discriminating and isolating segments and retrieving stored lexical forms.

In all three components, the Extended representations and processes are substantially more impaired than the Basic ones. Furthermore, the behavioral impairment with Extended grammar may be caused by abnormalities in the anatomy or functioning of distinct neural pathways, particularly dorsal circuits in which the arcuate fasciculus links regions in the superior posterior temporal lobe to the left inferior frontal gyrus, with further interconnections with the left basal ganglia. The pattern of impairment in G-SLI thus points to a set of neural pathways underlying Extended processing which appear to be a recently evolved, late developing, and genetically complex substrate for the most distinctively human cognitive trait (Box 2).

Box 2. Outstanding questions

- Can subtypes of SLI (such as the G-SLI discussed in this paper) be reliably identified across populations, languages, and laboratories?
- Can more sensitive methods of neural function, cortical and subcortical anatomy, and tractography identify the neural substrates of different language components more precisely?
- ‘Geno’s paradox’ is the failure to pinpoint specific genes that underlie massively heritable traits, presumably because most behavioral traits involve either many genes with small effects, rare genes with larger effects, or both. Will new genomic methods designed to resolve the paradox identify additional genetic contributors to SLI?
- What are the patterns of overlap and non-overlap in language disabilities, neuroanatomical differences, and genetic causes between different forms of SLI, other developmental disorders (e.g., autistic spectrum disorders), and degenerative disorders (such as Huntington’s and Parkinson’s disease)?
- Are there forms of language impairment that consist of impaired Basic representations and disrupted ventral pathways, but with relatively spared dorsal pathways?
- What are the clinical implications for focused remediation in children with SLI? Should we attempt to strengthen their Extended processes, or encourage them to compensate using their strengths in Basic representations?

Acknowledgments

This paper is a capstone to the research of Heather van der Lely, who died of cancer in February, 2014. She identified the grammatical subtype of SLI, characterized it empirically (with collaborators in London, Paris, Louvain la Neuve, and Berlin), proposed the Extended-Basic distinction to explain the pattern of deficits, and wrote the first draft of this paper. Steven Pinker rewrote and fact-checked the paper before and after Heather's death, with extensive input from her main collaborator Chloe Marshall and from John Harris, Gary Marcus, William Marslen-Wilson, David Poeppel, Friedemann Pulvermüller, Uli Sauerland, Michael Ullman, and two anonymous reviewers. This research was supported by the Wellcome Trust (grant 063713), the Economic and Social Research Council, the Medical Research Council, The EU 7th Programme (Grant CLAD 135295 LLP-1-2007-UK Kat SRC), the Leverhulme Trust, and the PD Fund at ZAS Berlin. Special thanks go to Heather's husband Paul Richens for drafting the figures, retrieving records, and much else.

References

- 1 Broca, P. (1861) Remarques sur le siege de la faculte du langage articule suivies d'une observation d'aphemie (parte de la parole). *Bull. Soc. Anat. Paris* 6, 330–357
- 2 Wernicke, C. (1874) *Der aphasische Symptomen complex*, Springer Verlag
- 3 Geschwind, N. (1979) Specializations of the human brain. *Sci. Am.* 241, 180–199
- 4 van der Lely, H.K. (2005) Domain-specific cognitive systems: insight from Grammatical-SLI. *Trends Cogn. Sci.* 9, 53–59
- 5 Walenski, M. et al. (2006) Language in autism. In *Understanding Autism: From Basic Neuroscience to Treatment* (Moldin, S.O. and Rubenstein, J.L., eds), pp. 175–203, Boca Raton, FL, Taylor and Francis Books
- 6 Cohen, N.J. et al. (2000) The interface between ADHD and language impairment: an examination of language, achievement, and cognitive processing. *J. Child Psychol. Psychiatry* 41, 353–362
- 7 Fisher, S.E. et al. (2003) Deciphering the genetic basis of speech and language disorders. *Annu. Rev. Neurosci.* 8, 8
- 8 Newbury, D.F. et al. (2010) Recent advances in the genetics of language impairment. *Genome Med.* 2, 6
- 9 van der Lely, H.K. and Marshall, C.R. (2010) Assessing component language deficits in the early detection of reading difficulty risk. *J. Learn. Disabil.* 43, 357–368
- 10 van der Lely, H.K. and Marshall, C. (2011) Grammatical-specific language impairment: A window onto domain specificity. In *The Handbook of Psycholinguistics and Cognitive Processes: Perspectives in Communication Disorders* (Gouendouzi, J. et al., eds), pp. 401–418, Psychology Press
- 11 van der Lely, H.K. and Stollwerck, L. (1996) A grammatical specific language impairment in children: an autosomal dominant inheritance? *Brain Lang.* 52, 484–504
- 12 Norbury, C. et al. (2002) Does impaired grammatical comprehension provide evidence for an innate grammar module? *Appl. Psycholing.* 23, 247–268
- 13 Rice, M. et al. (1998) Tense over time: the longitudinal course of tense acquisition in children with Specific Language Impairment. *J. Speech Lang. Hear. Res.* 41, 1412–1431
- 14 Leonard, L.B. (2007) Processing limitations and the grammatical profile of children with specific language impairment. *Adv. Child Dev. Behav.* 35, 139–171
- 15 Hamann, C. et al. (1998) German impaired grammar: the clause structure revisited. *Lang. Acquis.* 7, 193–245
- 16 Stavrakaki, S. and van der Lely, H. (2010) Production and comprehension of pronouns by Greek children with specific language impairment. *Br. J. Dev. Psychol.* 28, 189–216
- 17 Friedmann, N. and Novogrodsky, R. (2007) Is the movement deficit in syntactic SLI related to traces or to thematic role transfer? *Brain Lang.* 101, 50–63
- 18 Marcus, G.F. (2001) *The Algebraic Mind*, MIT Press
- 19 Pinker, S. (1999) *Words and Rules: The Ingredients of Language*, Weidenfeld & Nicolson
- 20 Chomsky, N. (1995) *The Minimalist Program*, MIT press
- 21 Fonteneau, E. and van der Lely, H.K. (2008) Electrical brain responses in language-impaired children reveal grammar-specific deficits. *PLoS ONE* 3, e1832
- 22 Adani, F. et al. (2013) Number dissimilarities facilitate the comprehension of relative clauses in children with (Grammatical) Specific Language Impairment. *J. Child Lang.* 41, 1–31
- 23 van der Lely, H.K. (1996) Specifically language impaired and normally developing children: verbal passive vs adjectival passive sentence interpretation. *Lingua* 98, 243–272
- 24 van der Lely, H.K. et al. (1998) Evidence for a grammar-specific deficit in children. *Curr. Biol.* 8, 1253–1258
- 25 Marinis, T. and van der Lely, H.K. (2007) On-line processing of wh-questions in children with G-SLI and typically developing children. *Int. J. Lang. Commun. Disord.* 42, 557–582
- 26 Pinker, S. and Ullman, M.T. (2002) The past and future of the past tense. *Trends Cogn. Sci.* 6, 456–463
- 27 Rice, M. et al. (2004) Grammatical tense deficits in children with SLI and nonspecific language impairment: relationships with nonverbal IQ over time. *J. Speech Lang. Hear. Res.* 47, 816–834
- 28 Ullman, M. and Gopnik, M. (1999) The production of inflectional morphology in hereditary specific language impairment. *Appl. Psycholinguist.* 20, 51–117
- 29 van der Lely, H.K. and Ullman, M.T. (2001) Past tense morphology in specifically language impaired children and normally developing children. *Lang. Cogn. Processes* 16, 113–336
- 30 Marshall, C.R. and van der Lely, H. (2012) Irregular past tense forms in English: how children with specific language impairment contribute to models of morphology. *Morphology* 22, 122–142
- 31 Marshall, C.R. and van der Lely, H.K. (2006) A challenge to current models of past tense inflection: the impact of phonotactics. *Cognition* 100, 302–320
- 32 Bozic, M. et al. (2013) Neurobiological systems for lexical representation and analysis in English. *J. Cogn. Neurosci.* 25, 1678–1691
- 33 Marshall, C.R. and Van der Lely, H.K.J. (2007) Derivational morphology in children with grammatical-specific language impairment. *Clin. Linguist. Phon.* 21, 71–91
- 34 Selkirk, E.O. (1984) *Phonology and Syntax: The Relation Between Sound and Structure*, MIT Press
- 35 Hickok, G. and Poeppel, D. (2007) The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402
- 36 Poeppel, D. et al. (2008) Speech perception at the interface of neurobiology and linguistics. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 363, 1071–1086
- 37 Ramus, F. et al. (2013) Phonological deficits in specific language impairment and developmental dyslexia: towards a multidimensional model. *Brain* 136, 630–645
- 38 Marshall, C.R. et al. (2011) Do children with dyslexia and/or specific language impairment compensate for place assimilation? Insight into phonological grammar and representations. *Cogn. Neuropsychol.* 27, 563–586
- 39 Marshall, C.R. and van der Lely, H.K. (2009) Effects of word position and stress on onset cluster production: evidence from typical development, SLI and dyslexia. *Language* 85, 39–57
- 40 Marshall, C.R. and van der Lely, H.K. (2007) The impact of phonological complexity on past tense inflection in children with Grammatical-SLI. *Adv. Speech Lang. Pathol.* 9, 191–203
- 41 Gallon, N. et al. (2007) Non-word repetition: An investigation of phonological complexity in children with Grammatical SLI. *Clin. Linguist. Phonet.* 21, 435–455
- 42 Marshall, C.R. et al. (2009) The link between prosody and language skills in children with specific language impairment (SLI) and/or dyslexia. *Int. J. Lang. Commun. Disord.* 44, 466–488
- 43 Friederici, A.D. (2009) Pathways to language: fiber tracts in the human brain. *Trends Cogn. Sci.* 13, 175–181
- 44 Friederici, A.D. (2012) The cortical language circuit: from auditory perception to sentence comprehension. *Trends Cogn. Sci.* 16, 262–268
- 45 Marslen-Wilson, W.D. and Tyler, L.K. (2007) Morphology, language and the brain: the decompositional substrate for language comprehension. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 362, 823–836
- 46 Shtyrov, Y. et al. (2003) Grammar processing outside the focus of attention: an MEG study. *J. Cogn. Neurosci.* 15, 1195–1206
- 47 Herrmann, B. et al. (2009) Localization of the syntactic mismatch negativity in the temporal cortex: an MEG study. *Neuroimage* 48, 590–600

- 48 Hanna, J. *et al.* (2013) Early activation of Broca's area in grammar processing as revealed by the syntactic mismatch negativity and distributed source analysis. *Cogn. Neurosci.* 5, 66–76
- 49 Friederici, A.D. (2012) Language development and the ontogeny of the dorsal pathway. *Front. Evol. Neurosci.* 4, 3
- 50 Oberecker, R. and Friederici, A.D. (2006) Syntactic event-related potential components in 24-month-olds' sentence comprehension. *Neuroreport* 17, 1017–1021
- 51 Friederici, A.D. (2002) Towards a neural basis of auditory sentence processing. *Trends Cogn. Sci.* 6, 78–84
- 52 Mestres-Misse, A. *et al.* (2012) An anterior-posterior gradient of cognitive control within the dorsomedial striatum. *Neuroimage* 62, 41–47
- 53 Ullman, M.T. *et al.* (1997) A neural dissociation within language: evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *J. Cogn. Neurosci.* 9, 266–276
- 54 Ullman, M.T. (2004) Contributions of memory circuits to language: the declarative/procedural model. *Cognition* 92, 231–270
- 55 Sahin, N.T. *et al.* (2006) Abstract grammatical processing of nouns and verbs in Broca's area: evidence from fMRI. *Cortex* 42, 540–562
- 56 Ullman, M.T. *et al.* (2014) Brain abnormalities in Specific language Impairment are localized to frontal regions and the caudate nucleus. *Under Revision*
- 57 Sahin, N.T. *et al.* (2009) Sequential processing of lexical, grammatical, and phonological information Within Broca's Area. *Science* 326, 445–449
- 58 Ullman, M.T., *et al.* (2014) The functional neuroanatomy of regular and irregular morphology. *In preparation*
- 59 Tyler, L.K. *et al.* (2005) Differentiating lexical form, meaning, and structure in the neural language system. *Proc. Natl. Acad. Sci. U.S.A.* 102, 8375–8380
- 60 Tyler, L.K. (2005) Temporal and frontal systems in speech comprehension: An fMRI study of past tense processing. *Neuropsychologia* 43, 1963–1974
- 61 Hestvik, A. *et al.* (2007) ERP measure of gap-filling in children with and without specific language impairment. In *Paper presented at the 20th CUNY Sentence Processing Conference*. UCSD, La Jolla
- 62 Sabisch, B. *et al.* (2006) Lexical-semantic processes in children with specific language impairment. *Neuroreport* 17, 1511–1514
- 63 Sabisch, B. *et al.* (2009) Children with specific language impairment: the role of prosodic processes in explaining difficulties in processing syntactic information. *Brain Res.* 1261, 37–44
- 64 Cantiani, C. *et al.* (2010) Impaired inflectional morphology in children with developmental dyslexia: converging evidence from behavioral and electrophysiological measures. In *BUCLD 36: Proceedings of the 36th Annual Boston University Conference on Language Development* (Biller, A.K., Chung, E.Y. and Kimball, A.E., eds), pp. 114–125, Medford, Mass., Cascadilla Press
- 65 Pizzoli, F. *et al.* (2007) Abnormal interaction between semantic and syntactic processing in specific language impairment: Electrophysiological evidence. In *Proceeding of the British Psychological Society* (Vol. 25), pp. 203–240, British Psychological Society
- 66 Shafer, V.L. *et al.* (2000) Deviant neurophysiological asymmetry in children with language impairment. *Neuroreport* 11, 3715–3718
- 67 Shafer, V.L. *et al.* (2001) Neurophysiological indices of language impairment in children. *Acta Otolaryngol.* 121, 297–300
- 68 Neville, H.J. *et al.* (1993) The neurobiology of sensory and language processing in language-impaired children. *J. Cogn. Neurosci.* 5, 235–253
- 69 Stromswold, K. (2001) The heritability of language: a review and meta-analysis of twin and adoption studies. *Language* 77, 647–723
- 70 Ganger, J. and Stromswold, K. (1998) Innateness, evolution, and genetics of language. *Hum. Biol.* 70, 199–213
- 71 Chabris, C.F. *et al.* (2013) Why is it hard to find genes that are associated with social science traits? *Am. J. Public Health* 103, 152–166
- 72 Enard, W. *et al.* (2002) Molecular evolution of *FOXP2*, a gene involved in speech and language. *Nature* 418, 869–872
- 73 Newbury, D.F. *et al.* (2009) *CMIP* and *ATP2C2* modulate phonological short-term memory in language impairment. *Am. J. Hum. Genet.* 85, 264–272
- 74 Whalley, H.C. *et al.* (2011) Genetic variation in *CNTNAP2* alters brain function during linguistic processing in healthy individuals. *Am. J. Med. Genet. B: Neuropsychiatr. Genet.* 156, 941–948
- 75 Newman, A.J. *et al.* (2011) The influence of language proficiency on lexical semantic processing in native and late learners of English. *J. Cogn. Neurosci.* 24, 1205–1223
- 76 Rice, M.L. *et al.* (2009) Convergent genetic linkage and associations to language, speech and reading measures in families of probands with Specific Language Impairment. *J. Neurodev. Disord.* 1, 264–282
- 77 Gardner, H. *et al.* (2006) Development of the Grammar and Phonology Screening (GAPS) test to assess key markers of specific language and literacy difficulties in young children. *Int. J. Lang. Commun. Disord.* 41, 513–540
- 78 van der Lely, H.K. *et al.* (2011) An investigation to validate the grammar and phonology screening (GAPS) test to identify children with specific language impairment. *PLoS ONE* 6, e22432
- 79 Jackendoff, R. (1997) *The architecture of the language faculty*, MIT Press
- 80 Stavrakaki, S. (2001) Comprehension of reversible relative clauses in specifically language impaired and normally developing Greek children. *Brain Lang.* 77, 419–431
- 81 Jakubowicz, C. and Tuller, L. (2008) Specific language impairment in French. In *Studies in French Applied Linguistics* (Ayoun, D., ed.), pp. 97–134, John Benjamins
- 82 van der Lely, H.K. and Battell, J. (2003) Wh-movement in children with grammatical SLI: a test of the RDDR hypothesis. *Language* 79, 153–181
- 83 van der Lely, H.K. *et al.* (2011) Who did Buzz see someone? Grammaticality judgement of wh-questions in typically developing children and children with Grammatical-SLI. *Lingua* 121, 408–422