Learning from Grammatical SLI

Response to J.B. Tomblin and J. Pandich (1999)

Heather K. J. van der Lely

I am indebted to Tomblin and Pandich for raising the issue of what can be learnt from children with specific language impairment (SLI) [Tomblin, J.B. and Pandich, J. (1999) Lessons from children with specific language impairment *Trends Cognit. Sci.* 3, 283–285][1]. My colleagues and I recently presented a detailed case study of a boy, AZ, who has an impaired grammatical system but intact cognitive and auditory systems[2]. While we did not suggest that all children with linguistic impairments (SLI) fit AZ's profile, we did show that six other children appear to show a similar profile, which I will refer to as Grammatical-SLI (G-SLI). We suggested that children with this profile support the existence of a genetically determined specialized mechanism that is necessary for the normal development of grammar. This suggestion has been challenged by Tomblin and Pandich on a number of grounds[1].

The domain specificity of the deficit

First, Tomblin and Pandich claim that vocabulary deficits that are within the language system but outside the grammatical system provide evidence against a specific and dedicated mechanism underlying grammar. Tomblin and Pandich interpret a high correlation between vocabulary and morphology scores as evidence to support the theory that the same mechanism underlies syntax learning and vocabulary development[3].

Although we agree that our children show a mild vocabulary impairment[2], an alternative is that there are (at least) two mechanisms involved in word learning[4] and the vocabulary impairment results from their grammatical deficit. One mechanism involves the use of perceptual and social cues, such as pragmatic inference, and is most relevant for learning concrete nouns (*horse, book*) and verbs (*cry, kick*)[4]. The other involves the use of structural-grammatical cues and is relevant for learning more abstract nouns (*idea, problem)* and verbs (*think, say*)[4,5]. Our children with G-SLI were normal with respect to logical reasoning and social-pragmatic inference[2] – the abilities likely to underlie the first mechanism for word learning outlined above. Consistent with this alternative, children with SLI are impaired in using grammatical cues to learn the meaning of novel verbs[6,7] and are particularly impaired in learning the semantic scope of quantifiers (*every, all*) (unpublished data), and abstract words and relational terms[8] – exactly those words and aspects of meaning for which grammatical cues are relevant.

In conclusion, vocabulary deficits might be predicted in children with a grammatical deficit because grammatical cues play an important role in word learning. I suggest that 'vocabulary' should not be thought of as a core, unified 'language system' *per se*; and that grammatical impairment causes the mild vocabulary and semantic deficits found in children with SLI.

Second, Tomblin and Pandich argue that the profile of AZ merely represents a selected extreme value in a normal distribution. Although, this is certainly a possibility, several considerations militate against this interpretation. First, AZ was just one of seven children who each showed the pattern of performance illustrated by AZ. Evidence suggests that they represent between 10–20% of children who have both persisting SLI and IQ abilities above 85 (Ref. 9). The chances of these seven children coming from the normal population are further diminished by several factors. First, the extreme grammatical standardized residual scores (SRs), which were based on the regression line for vocabulary, were often between –2 and –3 or below, but many of AZ's non-verbal cognitive SRs fell well above average. By chance alone, with SR scores of –2.0, –2.5, –3.0 and –3.9 you would expect just 2.3%, 0.6%, 0.14% and 0.0005%, respectively, of the normal population to have such deficits. Second, the findings were consistent across the range of methods used (spontaneous or elicited language, comprehension and judgement tasks) and for very different grammatical abilities, for example, assigning pronominal reference (*Mowgli says Baloo is tickling* ***him***), or affixing the ***-ed*** morphological tense marker. Third, our children show an atypical pattern of pragmatic-social and grammatical development that is inconsistent with Tomblin and Pandich's claim that they are merely 'inefficient learners'. In the story-telling task, our normally developing children under seven years incorrectly used pronouns to re-introduce a character but rarely made errors on the grammatical task involving pronouns, whereas the children with G-SLI show the reverse pattern of impairment[2,10,11].

The relationship between grammatical deficits and cognitive deficits

Tomblin and Pandich's suggestion that more general-purpose cognitive systems are important for language but are not restricted to language acquisition and use, is exactly what our data challenge because we found normal cognitive and auditory abilities alongside impaired grammatical abilities. We found no significant relationship on any of the many measures between cognitive abilities and grammatical abilities. From Tomblin and Pandich's perspective we might expect that a deficit in the cognitive system would reflect a more severe deficit such that, for example, children exhibiting cognitive deficits would omit a greater number of morphemes in their speech than those without cognitive deficits. Our data do not support this suggestion. Our children showed a severe deficit, failing on grammatical tasks that are normally mastered by children of five years old or younger, whereas children with Williams syndrome who had severe cognitive impairments performed faultlessly on three of our grammatical tests[12]. Finally, replications of our grammatical investigations with children with SLI who are not selected sub-groups but whose IQs are above 85, report remarkably similar patterns of impairment[7,13]. In addition, Bishop and colleagues' investigation of 144 twins with SLI revealed that approximately 10% (15) of the children scored highly on four or five qualitative markers of G-SLI, despite co-occurring cognitive deficits (pers. commun., D. Bishop *et al*.).

Tomblin and Pandich's claim for a relationship between cognitive and grammatical abilities is not strengthened if we compare only the measures of processing capacity (processing speed, short-term-memory or phonological capacity[8,14–17]) claimed to cause SLI. An investigation of lexical processing using an on-line auditory-auditory lexical decision task found that responses for 16 children with G-SLI were significantly faster than younger language-matched control children and, to our surprise, did not differ from age-matched control children[18]. Moreover, response speed did not correlate with grammatical abilities. An investigation of compound formation revealed that G-SLI children used regular plural nouns inside compounds (e.g. *rats-eater*), whilst normally developing control children rarely did[19]. This finding is not predicted by the processing account of SLI, but is consistent with a grammar-specific deficit in which regular plural forms are lexically stored[19].

In summary, we found normal cognitive abilities in seven children with G-SLI, a lack of any significant correlation on cognitive and grammatical performance, and results that conflict with the processing hypothesis when it is directly tested. These findings have to be 'worrying' at the very least for researchers who claim that grammatical deficits are caused by speed of processing or phonological memory deficits[8,15–17]. The evidence does not support a direct or indirect relationship between grammatical deficits and other cognitive deficits.

Tomblin and Pandich also object to our claim on the basis that particular

H.K.J. van der Lely is at the Department of Psychology, Birkbeck College, University of London, Malet Street, London, UK WC1E 7HX.

tel: +44 0207 631 6317
fax: +44 0207 631 6312
e-mail: h.vanderlely@bbk.ac.uk

1364-6613/99/$ – see front matter PII: S1364-6613(99)01354-6

grammatical deficits, such as tense, are not found cross-linguistically. First, this claim is controversial. The findings from cross-linguistic investigations of tense in French, Italian, Hebrew, Swedish and German[8,20–24] conflict and exemplify different linguistic characteristics for children with SLI. Second, in other recent work I have advocated that G-SLI is best characterized by a deficit in the computational-syntactic system[25]. Therefore, a more reliable cross-linguistic trait of G-SLI would be to look at aspects of grammar that reflect complex, dependent structural relationships, rather than one particular grammatical feature such as tense. Third, qualitative differences between children are also revealed within studies. An investigation by Neville, Tallal and colleagues revealed that a sub-group of children with SLI had impaired temporal auditory processing on behavioural and EEG measures, and another sub-group had impaired grammatical abilities on behavioural and EEG measures but normal auditory abilities[26].

These differences within and between studies of children with and without co-occurring cognitive deficits suggest differences in the underlying nature of the disorder. Due consideration of the similarities and qualitative differences in the forms of SLI and the nature of the linguistic deficit is required to understand fully the disorder.

Genes and grammatical deficits

Tomblin and Pandich highlight the general consensus that a genetic deficit causes SLI. At present we are far from understanding exactly how genes affect the development of neural pathways, to result in an impaired grammatical system, or SLI in general. The view one takes with respect to mechanisms versus representations and domain specificity, affects the inferences (or speculations!) one makes as to how genes cause SLI and the development of specialized cognitive systems. The distinction between mechanisms and representations is important and can be made theoretically[14]. In my view, Tomblin and Pandich make several incorrect inferences regarding this issue.

In our article we refer only to specialized mechanisms or cognitive systems and do not consider the issue of mechanism versus representation. We can speculate as to whether the deficit in the computational-syntactic system[25] is in a mechanism with a unique algorithm for grammatical processing, in the mechanism's 'attraction' to processing certain aspects of the input in certain ways, or in the capacity to form consistently particular structural representations given certain inputs. However, I know of no investigation into SLI where the data provide evidence distinguishing a mechanistic from a representational deficit. Logically, if a mechanism is impaired then this will impinge on the representations that develop from, or are served by, this mechanism. The distinction is a problem, which I acknowledge. But this issue should be separated from the questions of domain specificity and the identification of the genetic deficit(s) causing the various forms of SLI.

We proposed that the results for the seven children with G-SLI 'argue for the existence of a genetically determined specialisation of a sub-system in the brain required for grammar, and it appears for nothing else' (Ref. 2, p. 1257). Tomblin and Pandich's conclusion partly concurs with ours. They claim that there are likely to be genes that affect the brain systems or mechanisms that serve language. However, in contrast to our view, they consider that these genes will affect other (non-grammatical) functions and representations served by these brain systems, thus affecting a wide range of complex behaviours. The evidence from our data argues against this view. For example, why doesn't the genetic deficit cause cognitive deficits in our children? One possibility is that there is genetic variation in the disorder and this accounts for the heterogeneous nature of SLI. We need to know what (if any) are the genetic similarities and differences in the genetic deficit causing G-SLI and other forms of SLI, and whether unique genetic characteristics underlie G-SLI. As yet we simply do not have this information. What is clear is that if we do not investigate the genetic characteristics underlying pure forms of SLI such as G-SLI, in order to compare them with other forms of SLI, then we might lose a valuable source of insight into the role of genes in the development of specialized cognitive systems, and in particular of grammar.

Conclusion

The range of impaired grammatical abilities tested in a variety of ways, the pattern of development of grammatical and non-grammatical language abilities, and the extreme SR scores for seven children with a profile of G-SLI (Ref. 2) make it very unlikely that all seven children are from the normally developing population. The evidence does not support a direct or indirect relationship between the grammatical system and other core cognitive systems in children with SLI. This suggests that co-occurring cognitive and grammatical deficits have a subtly different origin, perhaps caused by genetic variation in the disorder. I propose that a primary, domain-specific grammatical deficit can cause secondary problems in learning words, thus impairing vocabulary development. In conclusion, I maintain that our results for the seven children with G-SLI argue for the existence of a genetically determined specialized mechanism required for normal development of grammar.

Acknowledgements

The author is supported by a Wellcome Trust Fellowship (044179/Z/95/Z). I thank Alastair McClelland for statistical advice and Gary Marcus for comments on an early draft.

References

1 Tomblin, J.B. and Pandich, J. (1999) What can we learn from children with exceptional language development? Lessons from children with specific language impairment *Trends Cognit. Sci.* 3, 283–285

2 van der Lely, H.K.J., Rosen, S. and McClelland, A. (1998) Evidence for a grammar-specific deficit in children *Curr. Biol.* 8, 1253–1258

3 Bates, E. and Goodman, J. (1997) On the inseparability of grammar and the lexicon *Lang. Cognit. Processes* 12, 507–584

4 Bloom, P. and Markson, L. (1998) Capacities underlying word learning *Trends Cognit. Sci.* 2, 39–76

5 Gleitman, L. (1990) The structural sources of verb meaning *Lang. Acquis.* 1, 3–55

6 van der Lely, H.K.J. (1994) Canonical linking rules: forward versus reverse linking in normally developing and specifically language impaired children *Cognition* 51, 29–72

7 O'Hara, M. and Johnston, J. (1997) Syntactic bootstrapping in children with specific language impairment *Eur. J. Disord. Commun.* 2, 189–205

8 Leonard, L. (1998) *Children with Specific Language Impairement*, MIT Press

9 van der Lely, H.K.J. and Stollwerck, L. (1996) A grammatical specific language impairment in children: an autosomal dominant inheritance? *Brain Lang.* 52, 484–504

10 van der Lely, H.K.J. (1997) Narrative discourse in Grammatical specific language impaired children: a modular language deficit? *J. Child Lang.* 24, 221–256

11 van der Lely, H.K.J. and Stollwerck, L. (1997) Binding theory and specifically language impaired children *Cognition* 62, 245–290

12 Clahsen, H. and Almazan, M. (1998) Syntax and morphology in children with Williams syndrome *Cognition* 68, 167–198

13 Precious, A. and Conti-Ramsden, G. (1988) Language-impaired children's comprehension of active versus passive sentences *Br. J. Disord. Commun.* 23, 229–244

14 Elman, J. *et al*. (1996) *Rethinking Innateness: A Connectionist Perspective on Development*, MIT Press

15 Joanisse, M. and Seidenberg, M. (1998) Specific language impairment: a deficit in grammar or processing? *Trends Cognit. Sci.* 2, 240–247

16 Tallal, P. *et al*. (1996) Language comprehension in language-learning impaired children improved with acoustically modified speech *Science* 217, 81–84

17 Gathercole, S. and Baddeley, A. (1990) Phonological memory deficits in language disordered children: is there a causal connection? *J. Mem. Lang.* 29, 336–360

18 Jones, M.J. and van der Lely, H.K.J. (1998) On-line lexical processing in specifically language impaired and normally developing children *Paper presented at the 23rd Boston University Conference on Language Development* (6–8 November 1998, Boston, MA, USA)

19 Van der Lely, H.K.J. and Christian, V. (1997) Lexical word formation in specifically language impaired children *Paper presented at the 22nd Boston University Conference on Language Development* (5–7 November, 1997, Boston, MA, USA)

20 Bottari, P., Chipriani, P. and Chilosi, A. (1996) Root infinitives in Italian SLI children, in *Proceedings of the 20th Boston University Conference on Language Development* (Stringfellow, A. *et al*., eds), pp. 76–86, Cascadilla Press, Boston
21 Dromi, E., Leonard, L. and Shteiman, M. (1993) The grammatical morphology of Hebrew-speaking children with specific language impairment *J. Speech Hear. Res.* 36, 760–771
22 Hakansson, G. (1998) Language impairment and the realization of finiteness, in *Proceedings of the 22nd Boston University Conference on Language Development* (Greenhill, A. *et al*., eds), pp. 314–324, Cascadilla Press, Boston
23 Hamann, C., Penner, Z. and Lindner, K. (1998) German impaired grammar: the clause structure revisited *Lang. Acquis.* 7, 193–245
24 Jakubowicz, C., Nash, L. and van der Velde, M. (1999) Inflection and past tense morphology in French SLI, in *Proceedings of the 22nd Boston University Conference on Language Development* (Greenhill, A. *et al*., eds), pp. 289–300, Cascadilla Press, Boston
25 van der Lely, H.K.J. (1998) SLI in children: movement, economy and deficits in the computational–syntactic system *Lang. Acquis.* 7, 161–192
26 Neville, H. *et al*. (1993) The neurobiology of sensory and language processing in language-impaired children *J. Cogn. Neurosci.* 5, 235–253

Correspondence

Networks are not 'hidden rules'

Does learning a language involve formulating rules or gathering statistics? Marcus and colleagues offered two pieces of evidence bearing on this debate[1]. Their behavioral studies were taken as evidence that babies form 'algebraic rules' and their attempt to model this behavior suggested that it is incompatible with the properties of a 'popular' class of connectionist networks. Both of these claims have been the subject of considerable discussion elsewhere[2–9]. In this letter we would like to raise a more general issue about the relationship between connectionist models and algebraic rules. More specifically, we would like to examine critically the statement by Marcus that 'Seidenberg and Elman have not gotten rid of the rule; they have simply hidden it', in reference to simulation data that we recently reported[2,5,7].

The purpose of our simulation was to demonstrate that the sequential regularities implicit in Marcus *et al.*'s stimuli provided a sufficient basis for differentiating sequences that conformed to the 'algebraic rule' from ones that did not, and that having acquired this information, a network could generalize appropriately to novel stimuli. The model was not an account of exactly how babies acquire this information; rather, it demonstrated that if they encoded this information, by whatever means, it would provide a basis for the observed behavior. That babies are able to detect such regularities is consistent with an extensive empirical literature[10].

Marcus objected to two aspects of this simulation: instead of the prediction task used in some other models (including Marcus *et al.*'s failed simulation), our model was trained to categorize stimuli as fitting a pattern or not doing so. We also used a supervised-learning procedure in which the network was provided with explicit feedback. In a sense, then, feedback to the model was structured on the basis of a rule; hence, Marcus concluded that the model must have had the rule 'hidden' in it. However, his statement is a *non sequitur*. Merely training a network to categorize stimuli into two groups using explicit feedback does not cause it to formulate a rule. Categorization, like prediction, is a task. The theoretical issue is how such tasks are performed. For a while now people have been debating two competing accounts. One is that such tasks involve formulating rules that have specific properties: for example, they operate over variables and are insensitive to incidental characteristics of the stimuli to which they apply. The other theory is different insofar as it states that the task is performed by using information that the rule mechanism explicitly ignores – stimulus properties, easily perceived by infants and learned by neural networks, that create different degrees of similarity between stimuli. Learning rules involves one type of procedure; acquiring statistical information involves another. Training a net to categorize stimuli using direct feedback does not cause one type of knowledge, a constraint-satisfaction network, to metamorphose into another type of knowledge, a rule, or create a rule-learning device out of one that performs gradient descent in weight space.

Marcus' concern was apparently prompted by the fact that the model was explicitly trained to turn a bit on for one type of sequence and off for another. But, the bit is not a variable and the network is not a rule. We choose this method as a simple and direct way to cause the model to encode statistical properties of the stimuli that we hypothesize underlie babies' categorization performance. As the Altmann[8] and Christiansen and Curtin[11,12] models show, there are other ways to achieve the same outcome, including ones more closely tied to the infant's experience (although see responses from Marcus[9,13]).

The network that we described, like other connectionist networks, provides a different account of behavior previously attributed to rules. The properties of such networks cannot be deduced from the properties of rules, at least as they are standardly construed. The only way to maintain the idea that the behavior of such systems is rule-governed is by changing the definition of 'rule' to mean 'having the properties of a neural network'. Stripped of its perjorative aspect, that is what the statement that the rule was 'hidden' in the network amounts to.

Pinker and Prince observed that in order to account for linguistic phenomena, connectionist networks would have to conform to ('implement') rule-based theories[14]. What has actually happened, as Marcus' comments illustrate, is that the concept of 'rule' is being altered to conform to the properties of connectionist networks. We think that this move involves abandoning what made rules an interesting alternative to connectionism. In any case, what is important is understanding the behaviour of the network and the principles that govern it, not Marcus' attempt to label it.

Mark S. Seidenberg
Neuroscience Program,
University of Southern California,
Los Angeles, CA 90089-2520, USA.
tel: +1 213 740 9174
fax: + 1 213 740 5687
e-mail: marks@gizmo.usc.edu

Jeffrey L. Elman
Department of Cognitive Science,
University of California,
San Diego, La Jolla,
CA 92093-0515, USA.
tel: +1 619 534 1147
fax: +1 619 534 678
e-mail: elman@crl.ucsd.edu

References
1 Marcus, G.F. *et al.* (1999) Rule learning by seven-month-old infants *Science* 283, 77–80
2 Seidenberg, M. and Elman, J.L. (1999) Do infants learn grammar with algebra or statistics? Letter *Science* 284, 434–435
3 Negishi, M. (1999) Do infants learn grammar with algebra or statistics? Letter *Science* 284, 435
4 Eimas, P.D. (1999) Do infants learn grammar with algebra or statistics? Letter *Science* 284, 435–436
5 Marcus, G.F. (1999) Do infants learn grammar with algebra or statistics? Response to Seidenberg & Elman, Negishi, and Eimas *Science* 284, 436–437
6 McClelland, J.L. and Plaut, D.C. (1999) Does generalization in infant learning implicate