

Reply to Garnham

ROBERT C. BERWICK
AMY S. WEINBERG*

Massachusetts Institute of Technology

Dr. Garnham (1983) seems to think that our (1983a) article is bad for psycholinguists—that we have done them a disservice by raising the Derivational Theory of Complexity (DTC) from its moribund slumber. Suprisingly, he takes this resuscitation to be the main point of our article:

Berwick and Weinberg draw the wrong conclusion from the fact that the DTC has not been refuted. They suggest that psycholinguists ought to give it a second chance. (p. 265)

Now, making psychologists ill does not seem like much of a research topic. Nor, to our mind, does the revival of the DTC: Dr. Garnham has missed our point. We came to bury (or at least expose the flaws of) the DTC, not to praise it. To review, we showed that if Transformational Grammar (TG) is embedded in an alternative parsing model, it can be made compatible with the reaction time results of Slobin (1966) and others. But such compatibility—at least with Slobin's results—is impossible if the underlying parsing model is assumed to operate in a strict serial fashion (as assumed by the DTC). Our conclusion was simply that, given the psycholinguistic evidence and a choice between changing either seriality or grammar, the assumption of seriality should give way. Transformational Grammar could be reconciled with the known experimental results.¹

Unfortunately, the demise of the DTC encouraged many psycholinguists to look elsewhere for knowledge representations appropriate for human parsing. Many approaches gave up the identification of the parser's knowledge representation with that representation designed to explain how children acquire language. In contrast, we believe that it is premature to abandon the

*Reprint requests should be sent to Robert Berwick NE43-814, M.I.T., Cambridge, MA 02139, U.S.A.

¹We thought this a valuable exercise because we believe that there is much to be gained by reincorporating TG into a theory of language processing. In Berwick and Weinberg (1983) we present positive proposals about how to join current TG theory to language processing—a combination that actually explains some outstanding puzzles in the theory of language comprehension.

strong hypothesis that one and the same knowledge representation drives both parsing and acquisition. We also felt that our paper would aid psycholinguists in the comparison of grammatical theories. This is simply because incompatibility with the DTC results has been used to motivate alternative grammatical theories—e.g., Lexical-functional (or Extended-Lexical theory) theory. Bresnan (1978) argued that a lexical-functional theory could be made compatible with the relevant DTC results and was therefore ‘better’ than a TG. Our paper suggests that parsing evidence of this type is neutral with respect to TG and LFG.

Dr. Garnham seems to think that we dissected the DTC in an *unilluminating* way, exposing its superficial injuries but missing the cause of the patient’s demise. In particular, he claims that showing TG to be incompatible with reaction time results is the inappropriate theoretical move. Rather, he insists, it is crucial to show that these types of experiments could not bear on the choice of possible grammars—that the DTC is *unfalsifiable* because any reaction time complexity attributed to a transformational component could be explained away by the contribution of other linguistic modules, e.g., word recognition or semantic interpretation:

In particular, the surface structure parser, the transformational component (if one exists), and the semantic interpreter are all potential sources of a difference in difficulty. (Garnham, 1983, p. 267)

Failing to specify possible complexity contributions of other components bears on the interpretation of the Slobin experiment—the main case we discussed—in the following way: Slobin found that subjects responded equally quickly to passive and active stimuli. However, a serially-embedded TG parser does not yield these predictions under the assumptions of the DTC because the passive transformation should take one more unit of time than a corresponding active. Garnham regards the failure to find a difference as ‘irrelevant’ because

... a difference in the transformational component could be offset by a difference in the reverse direction in some other component, though it is difficult to think of a plausible candidate in the case of actives and passives. (Garnham, p. 267)

But here we think Garnham is a bit ungenerous to earlier psycholinguists. He implies that the DTC experimenters were unaware that processing components could impose unequal burdens on comprehension tasks or that trade-offs in complexity could influence reaction times.

However, proponents of the DTC did not raise the issue of how long it took to compute a semantic interpretation. (Garnham, p. 267)

On this view, experimenters should have seen that their results were irrelevant to interpreting the complexity contribution of the transformational component.

We think though that the trade-offs Garnham proposes are implausible in the Slobin case. For this reason, we think that we were right to consider his results as potential problems, and to propose parsing mechanisms to deal with them. For example, on the assumptions of the standard theory, semantic interpretation works off of deep structure. Given the transformational model (the Standard Theory) then popular, actives and passives are identical at deep structure. The result: no difference in semantic interpretation for actives and passives, because semantic interpretation in both cases works off of an identical deep structure. Similar remarks apply to the suggested interference of a word recognition component. If word recognition were to compensate for the passive's excessive transformational complexity, then the offsetting factor would presumably be an extra cost associated with retrieving the active form of a predicate, compared with the retrieval of the passive form. But this cuts against the grain of lexical access experiments, where non-derived, more frequently occurring words (the active forms) should be retrieved more quickly (see Bradley, 1978; Forster and Chambers, 1973; Savin, 1963).

So, even taking into account the semantic interpretation and word recognition components, the Slobin results would still be counterexamples for a TG-based comprehension model. Earlier psycholinguists were right to worry; transformational grammar was compromised as the basis for the theory of language use, under the DTC assumptions. However, our paper details plausible alternative parsing assumptions that are highly compatible with the TG framework.

Taking a broader perspective where one is not so interested in the assumptions of earlier grammatical theories, one might still wonder what importance our article has for current work in psycholinguistics. In one sentence it is this: it is an attempt to work out some of the problems in mapping between competence and performance. In particular we show that neither a lexical-functional or an extended standard theory can be made compatible with reaction time experiments if we insist on their isomorphic representation in a language processing device. In Bresnan's (1978) system the lexical redundancy rules that handle the passive construction are non-isomorphically precomputed and stored as lexical templates. In the EST-based theory the passive transformation is replaced by operations compatible with a Marcus style parser (Marcus, 1980), in such a way that finite time is traded for finite space.

Considerations like homomorphic *versus* isomorphic embedding, serial *versus* parallel computational operations, and the like are relevant to evaluating the reaction time complexity of any component. So while we agree with

Garnham that it is necessary to keep trade-offs in mind, we claim that our paper is a logical first step toward understanding these trade-offs because it analyzes how any *one* component could influence processing complexity.²

Finally, Garnham seems to suggest that our discussion of Bresnan's interpretation of Slobin's data is less relevant than we would have hoped because we picked

... a set of rather tentative ideas proposed by Bresnan (1978), which I suspect she would no longer be interested in defending in detail, given more recent changes to her grammar, and her interest in chart parsing. (Garnham, p. 268)

Garnham should read Bresnan and Kaplan (1982, p. xxxvi ff.), where the ability of lexical-functional grammar to explain Slobin-type experimental results is defended in detail. In this section Bresnan and Kaplan also claim that our approach is still unable to square TG with possible reaction time results, given certain cases where transformations can be cascaded together to 'feed' the passive transformation. Transformational rules are in a 'feeding relationship' to one another when the structural description of one rule is not met until another rule has applied. The relevance of these remarks to our work is that parallel computation of rules in a feeding relationship is not possible because one rule must apply *before* the structural description of the other rule is met.

Here is what Bresnan and Kaplan say in part:

To successfully mimic the results of the LFG-based model however, they (Berwick and Weinberg rcb/asw) would have to demonstrate that the operations specified by *all* of the sequences of standard transformational operations ... can be executed in unit time. But because these operations are in true feeding relationships, in which the necessary input of one operation is created by the output of another, it is simply not possible to execute them in parallel. (Bresnan and Kaplan, 1982, pp. xxxv-xxxvi)

But this point is ill-founded, probably because it looks at the problem from the standpoint of generating sentences, rather than from the standpoint of language processing. 'Feeding relationships' that are established when mapping from deep to surface structure in the grammarian's derivation of sentences are not necessarily preserved when one maps from surface to deep structure (as one does in parsing). For example, consider the case of the

²We discussed the Slobin experiments because we felt that they highlighted these issues quite well. As we said in our main article (Berwick and Weinberg, 1983a, p. 10) we were primarily intrigued by the conceptual issues underpinning Slobin's work, not the actual results themselves. As we mentioned in the paper, we are aware that Forster and Olbrei have shown that Slobin's results may reflect inattention to the important psycholinguistic variable of *plausibility* and may thereby be compromised (see p. 10 of our article).

passive rule feeding the rule of There-Insertion. One shows that these rules are in a feeding relationship by showing that the structural description for 'There-Insertion' is not met until 'Passive' has applied. Assume that the structural description for 'There-Insertion' is something like the following:

(1a) Structural description: NP (aux) be X →
 1 2 3 4

(1b) Structural change: There be NP X
 2 3 1 4

Given a deep structure like (2a), we see that this structural description is not met until Passive has applied, yielding (2b).

(2a) e was being eaten an apple.

(2b) An apple_i was being eaten e_i

(2c) There was an apple being eaten.

In the corresponding *parse* however, we know that 'There-Insertion' has applied because we can recognize the lexical item 'THERE' which directly signals the application of this rule. Passive is also signalled independently by the passive morphology on the verb. 'There'-Insertion is marked by the presence of a specific lexical element. The relationship between it and the post-posed element is also a local dependency. Thus the parser can link this element directly, upon its recognition to the 'THERE' in the subject position. It does not have to wait until it has positive evidence that the passive rule has applied. This means, *contra* Bresnan and Kaplan that parallel processing of these two rules is still possible.³

In short, current LFG theory still strives to account for reaction time evidence like Slobin's. Further, the method chosen to obtain this compatibility is the same as that discussed in our earlier article. Exploring alternative processing designs and showing that Slobin-type evidence does not choose between competing linguistic theories is still then a valuable exercise.

Whatever the outcome of the psycholinguistic results, we took the aim of our article to be an exploration of the connection between computation and linguistic theory. Simply put, we tried to show how different grammatical representations could make different predictions regarding processing complexity—the first step in developing a grammatically based theory of language use.

³This case is actually quite representative. The other examples that Bresnan and Kaplan cite are subject to the same counter-argument or to the argument that Bresnan and Kaplan have misinterpreted what the transformational treatment of a given construction would be. For a full discussion see Berwick and Weinberg (1983b).

References

- Berwick, R.C., and Weinberg, A.S. (1983a) The role of grammars in models of language use. *Cog.*, 13, 1–61.
- Berwick, R.C., and Weinberg, A.S. (1983b) *The Grammatical Basis of Linguistic Performance*, Cambridge, MA, MIT Press.
- Bradley, D. (1978) Computational distinctions of vocabulary type. Unpublished MIT Ph.D. thesis.
- Bresnan, J., (1978) A realistic transformational grammar. In J. Bresnan, M. Halle and G. Miller (eds.) *Linguistic Theory and Psychological Reality*, Cambridge, MA, MIT Press.
- Bresnan, J. and Kaplan, R. (1982) Introduction. In J. Bresnan (ed.) *The Mental Representation of Grammatical Relations*, Cambridge, MA, MIT Press.
- Forster, K.I. and Chambers, S.M. (1973) Lexical access and naming time, *J. verb. Learn. verb. Behav.*, 12, 627–635.
- Garnham, A. (1983) Why psycholinguists don't care about DTC: A reply to Berwick and Weinberg. *Cog.*, 15, 263–270 (this issue).
- Marcus, M. (1980) *A Theory of Syntactic Recognition for Natural Language*, Cambridge, MA, MIT Press.
- Savin, H.B. (1963) Word frequency effect and errors in the perception of speech, *J. Acoust. Soc. Amer.*, 35, 200–206.