Note: This reply was submitted to Language in 1994, but was not accepted for publication.

An update version of this reply would include reference to Gibson's 1996 Cognition article that includes competition as a basic mechanism, a discussion of the new Frazier and Clifton "Construal" book that gives a larger role to pragmatic forces in parsing, and new mechanistic implementations of the Competition model.

The Competition Model and UG

Brian MacWhinney
Carnegie Mellon University

Elizabeth Bates
University of California at San Diego

In a recent issue of this journal, Edward Gibson published a review of 'The crosslinguistic study of sentence processing' (CSSP) -- a book which contains several chapters summarizing our recent crosslinguistic work and those of our collaborators in several countries. The primary goal of CSSP was to explore the predictions that derive from a minimalist model of language processing and acquisition that we have called the Competition Model.

Gibson's review of CSSP exposes a set of deep fracture lines separating two competing approaches to psycholinguistics. One one side of the fault zone, the Competition Model focuses on the construction of a minimalist model designed to predict exact numerical values in controlled studies. In this approach, an initial candidate model is not responsible for predicting all levels of all variables for all sentences in all languages. Instead, the model is developed inductively on the basis of a constrained set of sentence types and certain limited predictions. On the other side of the psycholinguistic fault zone, the UG model that Gibson espouses avoids minimalism and emphasizes the complexity of the theory of formal principles of language structure motivated by binary acceptability judgments generated by small numbers of professionally-trained theoreticians.

There are, in principle, three possible ways that one can understand the relation between the Competition Model and UG:

1. The Competition Model and UG could be viewed as non-competing accounts of totally separate aspects of reality. According to this analysis, comparing the Competition Model with UG would be like comparing programs for stock market technical analysis with models of bird flight. If two models describe different realities, they should be viewed as mutually irrelevant and should not be compared competitively. However, we believe that language is a unitary cognitive phenomenon. As a result, we cannot accept the idea that UG and the Competition Model are mutually irrelevant.

1

2. The Competition Model and UG could be viewed as non-competing and complementary accounts of different aspects of the same set of related phenomena. Although Gibson and others have rejected this possibility, we believe that it is premature to exclude this potential relation.
3. Finally, it could be the case that the two models are direct rivals and that one is right and the other wrong. This is the approach Gibson pursues. In particular, he chooses to argue that UG is 'adequate' and the Competition Model is 'fundamentally flawed.'

Gibson raises the following seven specific complaints:
1. The Competition Model fails to provide a role for the constraints proposed by UG.
2. The Competition Model fails to provide a constrained characterization of the core construct of a 'cue'.
3. The Competition Model fails to incorporate a parsing algorithm.
4. The Competition Model construct of 'conflict validity' is logically reducible to the construct of 'reliability' and hence vacuous.
5. Not enough work has been done on the empirical estimation of cue validities for texts and corpora.
6. The Competition Model does not offer a principled account of cue cost effects.
7. The use of ungrammatical stimuli in Competition Model experiments invalidates their results.

Some of these complaints are based on substantive disagreements about basic strategies in empirical research. Others are simple misunderstandings. Let us take a look at each of Gibson's complaints in turn.

<u>Complaint #1</u>. The Competition Model fails to provide a role for the constraints proposed by UG.

Gibson argues that constraints such as c-command (Reinhart, 1976) cannot be learned on the basis of positive evidence and, hence, must be innate components of UG. We are quite familiar with this argument and the literature based upon it. However, there are four reasons why we have not included a role for UG constraints in the current version of the Competition Model:

1. Arguments for constraints hinge primarily on the supposed non-availability of negative evidence. However, as MacWhinney (MacWhinney, 1987b; MacWhinney, 1987c; MacWhinney, 1988) and Bohannon, MacWhinney, and Snow (Bohannon, MacWhinney & Snow, 1990) have shown, learning within a Competition Model framework does not need to rely on negative evidence. The same type of blocking relation viewed as a 'benign' exception to the standard poverty of stimulus argument (Baker & McCarthy, 1981; Pinker, 1984; Pinker, 1989) is viewed in the Competition Model framework as operative throughout grammar. For example, the word 'release' is viewed as blocking the overgeneralization '*unsqueeze' (Bowerman, 1987; Bowerman, 1988) and the construction 'deliver X to Y' is seen as blocking the dative movement construction 'deliver XY'. Despite the fact that references to this work are given in Chapter 13 of CSSP, Gibson appears to have been unaware of the discussion of these issues by Bowerman, Bohannon et al., and MacWhinney. If he had been

aware of these analyses, he would have understood more clearly why the Competition Model provides no role for the constraints he is proposing.

2. There is little unanimity within the UG literature regarding the details of particular constraints. Even a constraint as well studied as c-command has no single received formulation (Kuno, 1986).

3. Applications of constraints to specific predictions in human sentence processing (Smyth, 1985) have cast further doubt on the usefulness of constraints in a processing model.

4. Most syntactic constraints do no apply to the sentences used in Competition Model experiments. Even the studies of complex relative clauses by MacWhinney and Pléh (MacWhinney & Pléh, 1988) and studies of anaphoric processes by MacDonald and MacWhinney (MacDonald & MacWhinney, 1990) and McDonald and MacWhinney (1995) have not needed to refer to constraints such a c-command.

We are not claiming that the phenomena underlying the various constraints described by GB theory are illusory or unimportant. Eventually, when the basis of these constraints in language processing and function is fully elucidated (Van Valin, 1990), these constraints will surely have an impact on the development of the Competition Model

Complaint #2. The Competition Model fails to provide a constrained characterization of the core construct of a 'cue'.

As noted in Chapter 13 of CSSP, this issue is discussed in detail in MacWhinney (MacWhinney, 1987b; MacWhinney, 1987c; MacWhinney, 1988). It is unfortunate that Gibson was unable to locate these chapters. They provide a detailed analysis in which syntactic role assignment is based on particular valence roles associated to particular lexical items. In those statements, the class of possible cues to valence roles is limited to: (1) the linear surface position of the argument *vis a vis* its head, (2) the affixes attached to the argument, (3) the affixes attached to the head, and (4) the inherent lexical semantic features of the argument. Unaware of these clearly stated limitations, Gibson portrays a search space full of 'monster cues' such as 'has an animate and a reptile to the left and a word that ends in a sibilant to its immediate right.' In fact the constraints already mentioned exclude this particular monster cue on at least three grounds. First, heads are single items and Gibson's cue requires the specification of two items to the left. Second, roles open slots for single items, but Gibson's monster cue tries to track not only two words to the left, but also the item to the immediate right. The Competition Model uses no such cues. Third, phonological forms are not potential cues to syntactic roles unless they have been analyzed as affixes. However, Gibson's monster cue is based on the tracking of a phonological form.

Gibson also worries that the cues discussed in CSSP require the induction of syntactic categories such as V or N and that no routine is provided for this inductive process. As explained in Chapter 13, the basic idea underlying the notion of 'preverbal' is that of 'linear position before the head', rather than 'linear position before V'. The fact that cues come to work similarly across large groups of lexical items is an emergent property of the connectionist architecture being assumed.

The emphasis on polysemic competition in lexical nets is a basic theme in Competition Model research (MacWhinney, 1982; MacWhinney, 1989). Gibson worries

that the Competition Model might confuse the role of 'the chicken' in 'the chicken cooked' with its role in 'the chicken cooked the dinner.' Gibson's fear is that all the model has to go on here is the cue of 'preverbal positioning.' If Gibson had been able to locate the references noted above (MacWhinney, 1987b; MacWhinney, 1987c; MacWhinney, 1988), he would have realized that the model postulates a competition in these cases between two polysemes of 'cook'. In these sentences, the presence of a postverbal noun would lead to the victory of the transitive polyseme and the absence of a postverbal noun would lead to the victory of the intransitive polyseme. Thus, comprehension is not solely dependent on the cue of 'preverbal positioning' but is a property of the overall system of competition, including both polysemic competition and role competition.

Gibson's various worries about monster cues betray not only his unfamiliarity with crucial aspects of Competition Model work, but also his failure to appreciate the power of connectionist models for problems involving cue extraction. For example, given an input vector of some hundred potential cues, it would be a trivial matter to train a back propagation network to extract a set of cues for predicting a handful of case roles (MacWhinney & Leinbach, 1991; MacWhinney, Leinbach, Taraban & McDonald, 1989; McClelland & Kawamoto, 1986; Miikkulainen & Dyer, 1991). If we expand the cue input space beyond 100, we might eventually choke up even a powerful recurrent network (Movellan & McClelland, 1992). But nothing so extreme is being envisioned here. Rather, the Competition Model only requires that the input include perhaps a dozen cues and the output track a handful of roles.

Complaint #3. The Competition Model fails to incorporate a parsing algorithm.

The series of articles that preceded publication of CSSP (MacWhinney, 1987b; MacWhinney, 1987c; MacWhinney, 1988) sketched out the shape of a Competition Model parser. The goal of this work was not to supplant current models in computational linguistics, but to explore ways in which competitive role processing can interact with a conventional parsing system. Kempen and Hoenkamp (Kempen & Hoenkamp, 1987) and Kempen and Vosse (Kempen & Vosse, 1989) have conducted similar explorations in the same spirit. Kempen's most recent model integrates a valence-based linear parser with a connectionist competitive role-assignment mechanism that uses simulated annealing. As in Kempen and Vosse, the focus of our work in on role competition and cue learning, rather than particular parsing algorithms, so we are happy to use conventional parsing algorithms until suitable connectionist models become available.

Complaint #4. The Competition Model construct of 'conflict validity' is logically reducible to the construct of 'reliability' and hence vacuous.

This complaint is based on a misunderstanding. The crucial difference between reliability and conflict validity is that reliability is computed on all sentences in which cue A appears. Conflict validity, on the other hand, in only computed on the based of those sentences in which cue A competes with cue B. Although it is true that all sentences involve potential conflicts between say cue A and some other cue we can call cue X, it is not true that all sentences necessarily involve conflicts between cue A and cue B. For example, the sentence 'him chased the bottle' has a conflict between the case cue that favors 'bottle' and the word order cue which favors 'him.' This particular conflict is

not present in a sentence such as 'The dog chased the bottle' where the case cue is neutralized. In terms of conditional probabilities, reliability is the probability that role assignment R is correct given cue A as a cue for R, or $p(R|A)$. The conflict validity of cue A is the probability that role assignment R is correct given the copresence of both cue A and the counterindicating cue B or $p(R||A^B)$. Note that $p(R|A) \neq p(R||A^B)$, unless we can guarantee that $p(A) = p(A^B)$ and this is seldom true.

Gibson's failure to understand this basic difference between reliability and conflict validity meant that he was unable to appreciate many of the central empirical contributions of the book, since he did not grasp the concepts involved.

Complaint #5. Not enough work has been done on the empirical estimation of cue validities for texts and corpora.

The establishment of tools for estimation of cue validities was the major driving force leading one of us (MacWhinney, 1991) to establish the CHILDES database and to push forward on the automatic morphological and syntactic analysis of that database. The computation of cue validities is a major task facing all psycholinguists and we are happy to get any help we can in advancing this important empirical work.

Complaint #6. The Competition Model does not offer a principled account of cue cost effects.

Gibson seems to misestimate the size of the contribution of cue cost factors to the current empirical database. The predictions made by the Competition Model have depended primarily on the basic mechanisms of reliability and conflict validity. This is particularly true for adult subjects. In only a few cases have we observed a misfit between cue strength and cue reliability during development. One such area is the late acquisition of verb agreement by Italian children. Another is the delay in the acquisition of use of accusative case marking by Hungarian children. We have treated these asynchronies not as glitches to be swept under the rug, but as major empirical findings that can be used to motivate a more complex theory of cue processing. These models are slowly converging with the CC-READER model of Just and Carpenter (Just & Carpenter, 1992) and even the account proposed by Gibson (1991). Undoubtedly, the many new on-line studies currently being conducted within the framework of the Competition Model (Kilborn, 1989; Li, Bates & MacWhinney, 1993; MacDonald & MacWhinney, 1990; McDonald & MacWhinney, 1995; Wulfeck, Bates & Capasso, 1991) will provide further stimulus to explicit integration of cue cost effects with the basic cue validity effects in a Competition Model parser.

Complaint #7. The use of ungrammatical stimuli in Competition Model experiments invalidates their results.

This complaint is certainly the one which we have heard most often from our colleagues. It is not surprising to find that linguists, who have been trained throughout their careers to focus on grammaticality judgments, should react sharply to grammaticality violations. However, we believe that there are five compelling reasons to include data from agrammatical sentences in our overall theory of language comprehension:

5

1. First, no sharp discontinuity has ever been observed in our studies between grammatical and ungrammatical sentences in terms of reaction times, choice percentages, or comprehension measures.
2. Moreover, in experiments in Croatian and Hungarian where a group of agrammatical stimuli happen to have the same syntactic form as a group of grammatical stimuli, the results for the two sets are nearly identical.
3. It should be remembered that students of perception (Cornsweet, 1970) have learned more from the study of illusions and paradoxical stimuli than from the study of everyday objects. Why should the study of speech perception be different?
4. What is important in these studies is not the individual results for an individual sentence, but the overall pattern of results and the extent to which it can be explained by a simple and coherent theory. When a large body of data shows a rich and orderly pattern, we should aim to understand these patterns, not to dismiss them (Stanovich, 1992).
5. Finally, there are many Competition Model experiments that use only grammatical stimuli. The model used to account for the results of these studies is the same one used to account for the results of studies that include some agrammatical stimuli.

To his credit, Gibson does understand the importance of the continuity demonstrations from Croatian and Hungarian noted in point #2. However he worries that, in these languages, the violations involve case marking, whereas the violations in English involve word order patterns. Relying on the notion of degrees of grammaticality (Chomsky, 1961), Gibson notes that '*him saw the dog' is not as bad as '*saw the dog the bird.' We agree. In English, word order is a strong cue and case is a weak cue. That is exactly the point of the Competition Model analysis. In Hungarian, the opposite is true and grammaticality judgments show the opposite pattern. This is precisely what the Competition Model predicts.

Some of Gibson's complaints are based on misunderstandings (#4) or incomplete readings of the Competition Model literature (#2 and #6). Some of his complaints have been effectively answered in the earlier published literature (#1, #2, and #7). Others (#3, #5, and #6) simply echo our own interests in continuing to deepen the empirical underpinnings of the Competition Model.   Gibson's review contributes no new insights regarding flaws in the Competition Model and his view of the Competition Model as 'fundamentally flawed' is not justified.

Gibson's analysis fails to come to grips with the empirical content of the research reported in CSSP. There is no discussion in his review of the language differences so fully documented in this book. There is no mention of the stability and replicability of the various cue interaction patterns, of the analyses of aphasic language (Bates, Wulfeck & MacWhinney, 1991), of the predictions regarding bilingual patterns (MacWhinney, 1987a), or of the investigations of learner varieties and the role of competition in sentence production. It is our hope that the lively debate between proponents of UG and supporters of the Competition Model does not obscure these basic facts of diversity in language processing. Instead, we hope that the debate has stimulated readers to pay increased attention to the actual empirical findings reported in CSSP.

REFERENCES

Baker, C. L., & McCarthy, J. J. (Eds.). (1981). The logical problem of language acquisition. Cambridge: MIT Press.

Bates, E., Wulfeck, B., & MacWhinney, B. (1991). Crosslinguistic research in aphasia: An overview. Brain and Language, 40, 1-15.

Bohannon, N., MacWhinney, B., & Snow, C. (1990). No negative evidence revisited: Beyond learnability or who has to prove what to whom. Developmental Psychology, 26, 221-226.

Bowerman, M. (1987). Commentary. In B. MacWhinney (Ed.), Mechanisms of language acquisition. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Bowerman, M. (1988). The "no negative evidence" problem. In J. Hawkins (Ed.), Explaining language universals. London: Blackwell.

Chomsky, N. (1961). Degrees of grammaticalness. Word, 17, 219-239.

Cornsweet, T. (1970). Visual perception. New York: Academic Press.

Just, M., & Carpenter, P. (1992). A capacity theory of comprehension: Individual differences in working memory. Psychological Review, 99, 122-149.

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. Cognitive Science, 11, 201-258.

Kempen, G., & Vosse, T. (1989). Incremental syntactic tree formation in human sentence processing: A cognitive architecture based on activation decay and simulated annealing. Connection Science, 1, 273-290.

Kilborn, K. (1989). Sentence processing in a second language: The timing of transfer. Language and Speech, 32, 1-23.

Kuno, S. (1986). Functional syntax. Chicago: Chicago University Press.

Li, P., Bates, E., & MacWhinney, B. (1993). Processing a language without inflections: A reaction time study of sentence interpretation in Chinese. Journal of Memory and Language, 32, 169-192.

MacDonald, M., & MacWhinney, B. (1990). Measuring inhibition and facilitation from pronouns. Journal of Memory and Language, 29, 469-492.

MacWhinney, B. (1982). Basic syntactic processes. In S. Kuczaj (Ed.), Language acquisition: vol 1. Syntax and semantics. pp. 73-136. Hillsdale, NJ: Lawrence Erlbaum.

MacWhinney, B. (1987a). Applying the competition model to bilingualism. Applied Psycholinguistics, 8, 315-327.

MacWhinney, B. (1987b). Competition and cooperation in language processing. In R. Tomlin (Ed.), Proceedings of the Pacific Conference on Linguistics. Eugene, Oregon: University of Oregon.

MacWhinney, B. (1987c). The Competition Model. In B. MacWhinney (Ed.), Mechanisms of language acquisition. pp. 249-308. Hillsdale, NJ: Lawrence Erlbaum.

MacWhinney, B. (1988). Competition and teachability. In R. Schiefelbusch & M. Rice (Eds.), The teachability of language. pp. 63-104. New York: Cambridge University Press.

MacWhinney, B. (1989). Competition and lexical categorization. In R. Corrigan, F. Eckman, & M. Noonan (Eds.), Linguistic categorization. pp. 195-242. New York: Benjamins.

MacWhinney, B. (1991). The CHILDES project: Tools for analyzing talk. Hillsdale, NJ: Erlbaum.

MacWhinney, B., & Leinbach, J. (1991). Implementations are not conceptualizations: Revising the verb learning model. Cognition, 29, 121-157.

MacWhinney, B., & Pléh, C. (1988). The processing of restrictive relative clauses in Hungarian. Cognition, 29, 95-141.

MacWhinney, B. J., Leinbach, J., Taraban, R., & McDonald, J. L. (1989). Language learning: Cues or rules? Journal of Memory and Language, 28, 255-277.

McClelland, J. L., & Kawamoto, A. (1986). Mechanisms of sentence processing: Assigning roles to constituents. In J. L. McClelland & D. E. Rumelhart (Eds.), Parallel Distributed Processing. Cambridge, MA: MIT Press.

McDonald, J. L., & MacWhinney, B. J. (1995). The time course of anaphor resolution: Effects of implicit verb causality and gender. Journal of Memory and Language, 34, 543-566.

Miikkulainen, R., & Dyer, M. (1991). Natural language processing with modular neural networks and distributed lexicon. Cognitive Science, 15, 343-399.

Movellan, J., & McClelland, J. (1992). Three learning rules for stochastic neural networks. Neural Computation.

Pinker, S. (1984). Language learnability and language development. Cambridge, Mass: Harvard University Press.

Pinker, S. (1989). Learnability and cognition: the acquisition of argument structure. Cambridge: MIT Press.

Reinhart, T. (1976). The syntactic domain of anaphora: M. I. T.

Smyth, R. (1985). Cognitive aspects of anaphora judgment and resolution. Bloomington, IN: Indiana University Linguistics Club.

Stanovich, K. (1992). How to think straight about Psychology. (Third Edition ed.). New York: Harper Collins.

Van Valin, R. (1990). Functionalism, anaphora, and syntax. Studies in Language, 14, 169-219.

Wulfeck, B., Bates, E., & Capasso, R. (1991). A crosslinguistic study of grammaticality judgments in Broca's aphasia. Brain and Language, 41, 55-72.

Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213