Gradience of Gradience: A reply to Jackendoff

JAMES L. MCCLELLAND AND JOAN BYBEE

Abstract

Jackendoff and other linguists have acknowledged that there is gradience in language but have tended to treat gradient phenomena as separate from the core of language, which is viewed as fully productive and compositional. This perspective suffuses Jackendoff's (2007) response to our position paper (Bybee and McClelland 2005). We argue that gradience is an inherent feature of language representation, processing, and learning, and that natural language exhibits all degrees of gradience. Contrary to Jackendoff's assertions, we do not reject the possibility of innate constraints on language, feeling only that the jury is out on the nature and specificity of such constraints. We address a number of questions Jackendoff raises about the process of grammaticalization, drawing on extant literature of which he appears to be unaware. We also address Jackendoff's views on the prospect that connectionist models can address core aspects of language processing and representation. Here again extant literature of which Jackendoff seems unaware addresses all four of his general objections to connectionist approaches.

1. Introduction

As might be expected between researchers with very different assumptions and goals, there seems to be a failure of communication between us and Jackendoff. He suggests that we have misunderstood his position, and, likewise, it seems to us he misunderstands ours. Yet, one point he makes is that we are not really as far apart in our views as one would think. In the interest of achieving greater understanding and in hopes of an eventual convergence, we respond here to several of his points.

First, we acknowledge that our paper took as its point of comparison a more Chomskyan theory of language, somewhat different from Jackendoff's position

The Linguistic Review 24 (2007), 437–455
DOI 10.1515/TLR.2007.019

0167–6318/07/024-0437
©Walter de Gruyter

as expressed in *Foundations of Language* (FL; Jackendoff 2002). We recognize that he embraces a number of properties of grammars, such as gradience in certain domains, that earlier generative theory would not. This is indeed something we share with him as a reaction against the strictest form of the Chomskyan framework. What we would add, however, is that Jackendoff may not have allowed his position to move far enough in a new direction – one reflected in the approaches we have attempted to represent. What we write below is intended to make our approach as clear as possible, since it seems apparent it is sometimes misunderstood. We address four areas of misunderstanding of our position that were brought up in Jackendoff's discussion: productivity, compositionality, innateness and grammaticalization. In some cases we feel that Jackendoff has misconstrued our position; in others, he appears not to appreciate some findings and developments in the experimental and computational literatures. In this response we will try to set the record straight concerning these issues. We then address Jackendoff's views on the prospect of connectionist models to address basic aspects of language processing and representation. Here again misunderstanding is rampant – there is a large extant literature of which Jackendoff seems unaware that addresses all four of his key objections.

The first two issues we will address concern productivity and compositionality in language, areas in which we find ourselves in clear disagreement with Jackendoff's perspective. To us the evidence indicates that productivity and compositionality in phonology, morphosyntax and semantics are expressed in language as matters of degree rather than as discrete categories. In fact, there is a full continuous range of both productivity and compositionality across phenomena (inflection is more productive and compositional than derivation) and within phenomena (irregular past tense is less productive and compositional than regular past tense), thus a gradience of gradience. The continuous nature of these dimensions and the dynamic nature of language, which is always changing, points clearly to linguistic structure as always emerging during language use (see Hopper 1987; Bybee and Hopper 2001), and to representations that are continuously shaped in response to the structure in experience (Bybee 1985, 2006; Rumelhart, McClelland et al. 1986; McClelland, Rumelhart et al. 1986).

2. Productivity

Let us first consider productivity. Jackendoff (2007: 377) says, discussing his own view of acquisition, 'Under this account, an important problem for acquisition is to determine which linguistic phenomena are productive and which are not' and then adds a footnote saying 'For Bybee and McClelland, this problem doesn't arise, of course, because they believe there are no truly productive phenomena.' Jackendoff is correct that the problem does not arise, but not for the reason stated.

First, lest Jackendoff or anyone else has the impression that we are uninterested in the fact that language is productive, we note that we have both taken productivity as an important research focus (Bybee and Slobin 1982; Bybee and Moder 1983; Bybee 1985, 1995; Bybee and Eddington 2006; Rumelhart and McClelland 1986). The goals of the Bybee work here cited were to determine exactly how productivity arises, and how it is maintained and acquired. The importance of type frequency was emphasized in this research as well as the interaction of token and type frequency. Degrees of productivity have been recognized as resulting from similarity of patterns (phonological and semantic) in interaction with type frequency. In our TLR paper (Bybee and McClelland 2005) we also cited research by Hay (2001, 2002), who has examined a variety of factors that determine productivity. Rumelhart and McClelland (1986) proposed their past-tense model in part to address the hallmarks of productivity in the regular past tense, namely use of the *+ed* past tense marker on nonce forms and in over-regularizations. So, clearly the issue of productivity and what determines its degree is a very important issue for us.

What about the statement that we 'believe there are no truly productive phenomena'? We would not state our position this way. To us, an aspect of language is truly productive if it exhibits generalization from experienced forms to novel forms. What we would say instead is that *we believe there is no dichotomous distinction between productive and unproductive phenomena, rather, there are only degrees of productivity*. This view – that there is no such distinction – was the heart of the claim in Rumelhart and McClelland (1986) that a connectionist network could account for production both of exceptions and of regular forms, and for the existence of tendencies in production to regularize as well as irregularize, as documented in Bybee and Slobin (1982) (and, indeed, in the earlier study by Berko 1958). One advantage of our proposal is that a single system exhibiting a range of productivity need not face the Gordian knot of dividing phenomena into only those that are productive and those that are not.

In FL, Jackendoff defines productive processes as those governed by rules containing a variable – that is, symbolic rules (2002:187). Our approaches (either the connectionist approach or an exemplar plus associative network approach as in Bybee 1985, 1995, 2001) do not rely on symbolic rules, but rather postulate that productive patterns are built up from experience with exemplars of multiple types. We define productivity as the tendency for a pattern to apply to new forms, a definition that allows productivity to be a matter of degree. There are many degrees of productivity, depending upon type frequency and other factors, as mentioned above.

When Jackendoff says that we believe there are no truly productive phenomena, he is confusing a belief about the underlying mechanism with a belief about the empirical phenomena. Though not the same in all respects, his proposals are similar to Pinker's (1991) dual-mechanism approach in that they make what we believe to be artificial distinctions like 'semi-productive' vs. 'truly productive'. The evidence suggests to us that productivity forms a continuum; nothing in our theories rules out highly productive, even exceptionless, patterns. However, we also maintain that the determination of productivity is an empirical matter and should be handled with care. Some highly productive patterns, including some morphological patterns and many constructions, can have large ranges of application. Such a large range would resemble a category represented by a variable. But in fact, as Croft (2001) has shown in considerable detail for passive and other constructions across languages, the construction itself defines the class and there are no variable-like classes that are independent of the constructions in which they occur.

A further, and crucial point, is that a tendency toward productivity is inherent in the approaches that we have advocated. We have offered them as alternatives to rule-based accounts of productivity. It is true that initial connectionist models, both of past tense formation and of single-word reading, that were based on these ideas (Rumelhart and McClelland 1986; Seidenberg and McClelland 1989) were not as productive as human language users. This left some room for critics to respond with the assertion that symbolic rules were needed to rise to the level of human performance. But in both of these domains, subsequent work has produced connectionist models that exhibit a degree of productivity matching the degree of productivity found in empirical studies with human subjects (Plaut et al. 1996; MacWhinney and Leinbach 1991). We consider connectionist and related approaches to sentence-level productivity below.

In FL, Jackendoff presents no evidence for a discrete division between semi-productive and fully productive patterns. He simply asserts that if rules don't have variables then all new formations – new instances of constructions – have to be created by analogy to existing stored instances (2002: 188), implying that analogy as a mechanism is inadequate to account for the full range of productivity that is seen in natural languages. We have two responses: First, while connectionist models share features with analogical models, they differ from them in many ways; and thus it is incorrect to partition the possible mechanisms that may underlie productivity in this way. Second, there are analogical models (Skousen 1989, 1992) that have been shown to capture the degrees of productivity found in natural language, including the highest degree of productivity.

In fact, as Jackendoff continues his discussion (2002: 189) he seems to acknowledge that productivity in language is always a graded phenomenon. He says: "Actually, the existence of an l-rule in the brain is more likely a matter of degree. The ease or speed with which an l-rule is activated relative to stored forms undoubtedly plays a role in how freely productive it is in performance." This step – assigning a graded strength to a categorical rule – is one that other proponents of symbolic rules, including Pinker (c.f. Pinker and Ullman 2002) take as well to salvage the two-process approach in the face of evidence of gradedness even among forms with a high degree of productivity. To us, however, this move undercuts both the empirical and the theoretical base on which an appeal to the existence of such rules is built, not only by Jackendoff but also by Pinker and by Marcus (2001). Both the phenomenon and the mechanism are no longer categorical in nature, as asserted in the very framing of the supposed dichotomy between truly and partially productive processes. Worse still, this move is not adequate, since the evidence of gradedness includes graded degrees of applicability as a function of semantic and phonological variables (Albright and Hayes 2003; McClelland and Patterson 2002). Rules that apply to varying degrees depending on the particular details of the filler of a variable slot can certainly be written, but such rules violate the 'algebraic' principle articulated with such insistence by, especially, Marcus (2001).

In summary, Jackendoff's position appears to contain within it the seed of its own destruction – namely, the belief that productivity in language is indeed always a matter of degree. We believe the attempt to maintain a dichotomy will not survive for long.

3. Compositionality

Jackendoff also claims that Bybee and McClelland reject 'the necessity for combinatorial mechanisms' (2007: 352). Again, Jackendoff mistakes the hypothesis that symbolic rules are not needed for the claim that there are no combinatorial mechanisms at all. Jackendoff's claim in FL (Chapter 3) that compositionality requires symbolic rules and constraints with algebra-like variables is just a claim – one that depends on the unsupported belief that compositionality is categorical in nature. As with productivity, we view compositionality as a continuum. The evidence cited in our paper is synchronic, diachronic and experimental. There is now considerable literature finding that usage factors such as relative and absolute frequency affect compositionality (Bybee and Scheibman 1999; Haiman 1994; Hay 2001). Such continua again support the idea of a single mechanism for dealing with the whole range of phenomena.

Graded compositionality is not just a matter of form, but is also a property of the semantics and the degree of compositionality of the form and meaning tend to be parallel (Bybee 1985). Consider, for example, the distinction between the occurrence of an event in the past and the occurrence of an event in the future. Because the present is ever changing, but the world is not, the

kinds of events that could have occurred yesterday and the kinds of events that could occur tomorrow are very similar – indeed all of the events that will occur tomorrow will be events of the past just two days hence. Thus, there is a high degree of independence (please notice that we do not say there is absolute independence) between past vs. future occurrence on the one hand and other aspects of the content of events on the other. Bybee (1985) suggested that this independence contributes to the properties that tense has as an inflectional category, and Lupyan and McClelland (2003) show that a simple connectionist network tends to maintain and even promote grammaticalization more strongly when the semantic representation of an intended utterance about an action is compositional than when it is not. Note that the compositionality is part of the cognitive representation of the event on this view, and is not specifically a linguistic matter at all. The consequence of this cognitive compositionality is a high degree of compositionality in both semantics and formal expression in the resulting language.

4. Innateness

We were particularly puzzled by the following statement by Jackendoff (2007: 371):

> Bybee and McClelland are rejecting the possibility of *any* innately structured behavior, from sneezing to birdsong to sex to parental care to complex social interaction, in disregard of the voluminous animal literature (e.g., Wilson 1980; Gallistel 1990; Hauser 2000).

We are mystified because we did not mention innateness in our paper at all. Nor did we mention anything about non-human species. We certainly never mentioned sneezing or sex! Perhaps our emphasis on learning, domain-generality, and extraction of patterns from experience makes it seem as though we eschew all possible forms of innate bias or pre-disposition. To set the record straight, we have no doubt at all that the genome endows organisms with tendencies that serve their own continuation in their ecological niche, and that the human genome in particular endows human organisms with tendencies relevant to the uniquely human ability to make use of language. Such tendencies may be quantitatively different from similar tendencies in related species, but they are not necessarily qualitatively different (Elman 2005). Connectionists have emphasized architectural constraints – e.g., network topology and constraints on connectivity – but there is also good reason to believe there are specialized neuronal- and circuit-level properties in different brain regions that make one brain area better suited to one type of task than another. We do not doubt some such properties arise in part through evolutionary selection pressure favoring

properties useful for language and communication. For example, the frequency sensitivity of the human ear is likely to match the frequencies effectively produced by the human vocal apparatus, and this sensitivity is likely to have been shaped by evolutionary selection.

It is true that we tend to disfavor the search for innate linguistic universals, at least universals expressed directly in the form of constraints on grammar as such, and we do emphasize what is general in cognition that may also apply to language. In any case, the evidence base for the existence of language universals is thin at best. Generative linguists tend to look for universals in one or two languages, since they believe languages are based on innate structures (see Newmeyer 2005 for a critique of this research program). More empirical approaches using samples of fifty, seventy-five or hundreds of languages provide a more valid basis for making claims about Universal Grammar. Yet researchers who do investigate large samples find many similarities among languages, especially in the way they change over time, but very few absolute synchronic universals of linguistic structure (see Bybee 1985, for example).

A further point is that something that generally holds across languages need not be innate or specific to language. For example, Jackendoff proposes a 'piece' of Universal Grammar that says that prototypically a phonological word corresponds to a noun, verb, adjective or preposition. A second 'piece' of UG says that a prototypical NP is a physical object and a prototypical VP is an action. While these may be fairly universal, it is not clear at all that there is a need to invoke innateness of features specific to language, since the relevant prototypes may well be learnable from shared features of human experience. Such features include, for example, the fact that humans experience the world as containing actors, actions, and objects and that there is a degree of independence between particular actors/objects and the actions they can perform or have performed on them. In addition, we wonder what it means for a so-called innate linguistic universal to be not absolute, but rather to identify only a prototype. It could mean that the observed tendency is the result of the convergence of various factors rather than being innately specified as such (again, see Newmeyer 2005).

Goldin-Meadow (2005) presents evidence that homesigners use predicate frames, that is, that they know how many entities are involved in an event. Children could learn these event structures from experience or they could be innate. However, what is not clear is whether they are strictly linguistic or a part of general cognition. Calling them 'predicate frames' makes them sound linguistic. Calling them 'event structures' (Croft 1991) or Conceptual Schemas (Slobin 1985) places them in a more general cognitive framework.

As is well-known, innateness is often invoked in linguistics as a form of explanation. Different researchers have different preferences with regard to how they formulate explanation and where they look for it. We prefer to consider

other possibilities for explanation before turning to innate specification. Similarly we prefer to examine domain general principles before turning to domain specific ones. We also prefer models that aim to capture the acquisition and use of language directly, rather than leaving these separate from an abstract specification of competence. Jackendoff's comments have not convinced us that we should turn away from these preferences. Indeed, they persuade us that his approach would be enriched by joining us in a search for accounts based in general-purpose, experience-dependent mechanisms.

5. Grammaticalization

In our introductory section we mentioned the study of grammaticalization as a major development that accounts for a major share of the gradience in grammar and also serves as an important source of evidence for the cognitive representation of grammar. Jackendoff recognizes the potential importance of grammaticalization, but unfortunately has not familiarized himself with the extensive literature on this topic, where great strides have been made in the last twenty years. Jackendoff (2007: 372) says:

> Although it is undeniable that grammaticalization through historical change is an important source of closed-class morphology, I find a lot left unexplained.

The questions he asks about grammaticalization are indeed addressed in the literature and much empirical evidence has been brought to bear on their answers. The result of this work presents an interesting and highly informative view of grammar as an outcome of language use. Two general overviews are available in Bybee (1998), which is particularly relevant to issues in the evolution of grammar, and Bybee (2003a), which is aimed at a general academic audience. Consider now how the questions Jackendoff (2007: 372) asks have been addressed.

> What is it about the semantics of *go*+purpose that lends itself to being bleached out into a future? And why can something that means 'want' bleach out to practically the same thing (English *will*)?

First, it is important to note that the word 'bleach' is not quite right in this context. As Traugott has shown in her many publications on this issue, semantic bleaching is only one aspect of the semantic change that occurs in grammaticalization. The other important mechanism is pragmatic strengthening or the conventionalization of implicature (Traugott 1982, 1989, 1995; Traugott and Dasher 2002). Frequently made inferences come to be part of the meaning of an expression, increasing its distribution and weakening its earlier, more lexical senses. In the case of the development of future markers, it has been

documented that expressions of movement towards a goal, volition and obligation can in context take on the inference of intention by the subject (Bybee and Pagliuca 1987; Bybee, Pagliuca and Perkins 1991; Bybee, Perkins and Pagliuca 1994). In fact, even temporal adverbials, which are much less frequently the source of futures, take on an intention reading on their way to becoming futures (Romaine 1995). A further inference from intention (especially about a third person subject) is one of speaker's prediction. That is, if I say that someone is intending to do something, the hearer can infer than I am predicting s/he will do it. In our semantic theory, prediction is the basic sense of future. Thus the convergence of expression from various sources into future is due to the inference that produces the intention meaning, and a subsequent inference that produces the prediction meaning. The modal nuances that are often found cross-linguistically to be expressed by future morphemes are either retentions from earlier more lexical meanings (such as volition and obligation) or extensions from prediction (see Bybee and Dahl 1989; Bybee, Perkins and Pagliuca 1994).

Next, Jackendoff (2007: 372) asks:

> More generally, what explains the "possible trajectories of change" (2005: 386) in the semantic domain? For instance, why couldn't *going to* reduce phonologically to *gonna* without bleaching its original meaning?

These questions seem to be (i) why are there unidirectional paths of change and (ii) why do form and meaning change together, rather than only form changing? These are fascinating and difficult questions, and they have been addressed in the extensive grammaticalization literature, which is as available to Jackendoff as to anyone. It is not possible to do justice to the first question in this short reply, as the answer is complex and open to some differences of interpretation. The basic answer is that although the paths of change found in grammaticalization are cross-linguistically similar, the more basic universals are the mechanisms of change that are operative in all languages at all times whether they be the use of metaphor, pragmatic inference, or generalization of meaning (Heine and Reh 1984; Heine, Claudi and Hünnemeyer 1991; Bybee, Perkins and Pagliuca 1994; Traugott 1982, 1989). In addition to the mechanisms of semantic change just mentioned, processes affecting form occur: phonetic reduction, reduction in compositionality, and change in constituent structure. Interestingly, similar lexical items enter into grammaticalizing constructions across languages, pointing to a cognitive basis for the changes, as indeed, the important role of inference also points to a cognitive-interactional basis for the semantic changes documented. If we think of language as a complex, self-organizing system with a relatively small number of processes that act on certain semantic and formal linguistic material, then we can see grammar evolving over time in a limited number of ways.

The second part of Jackendoff's question, why phonetic reduction couldn't take place independently of semantic change, is much less complex and therefore somewhat easier to answer. First, observe that phonetic reduction does take place independently of semantic change in some cases. One might cite the fusion of prepositions and articles in languages such as Spanish, where *a + el* becomes *al* or similar changes in French and German. Frequently-used sequences tend to undergo fusion and reduction even if there is no special semantic affinity among the elements or any semantic change going on. Another example is the contraction of the English auxiliary with subject pronouns, as in *I'm* and *I'll*. However, in the case of *be going to*, the requisite frequency of use is made possible by the pragmatic and semantic change which increases the contexts of use for the expression and thereby increase its frequency. Then it is the frequency of use that conditions the extreme phonetic reduction. Thus the semantic change is paralleled by the phonetic reduction (Bybee 2003b).

Jackendoff's (2007: 372) next question has also been examined intensely in the literature:

> Why do certain sorts of meanings get encoded as closed-class morphemes while others don't?

This is indeed a basic question about human language. In the generative paradigm it can only be answered by saying that the meanings of closed-class morphemes are innate (Cinque 1999; Roberts and Roussou 2003). Presumably Jackendoff is not satisfied with this answer, as it cuts off further enquiry and says nothing about the complex interaction between cognition, function in discourse, and evolutionary development. Cognitive-functional linguists have not been satisfied with this type of answer either, and have approached the question of which meanings are expressed lexically and which grammatically from a cognitive point of view (Talmy 1985, 1988). Bybee (1985) discusses both semantic factors and distributional factors in explaining the meanings and positioning of verbal affixes. Slobin (1997) argues against the position that children have an innate catalog of meanings that can be grammaticized, noting for instance that the content of some grammaticalized categories, such as epistemic modals, which express possibility and probability, is difficult for young children and only acquired after children learn more about the pragmatics of conversation. The other evidence against the innateness of grammatical notions is that they evolve gradually from lexical items and differ in detail and nuance across languages. The question of why languages develop very similar futures, pasts, progressives, perfects, perfectives, imperfectives, etc. is addressed in the last chapter of Bybee, Perkins and Pagliuca (1994). The evidence from the way such categories are used at different stages of development demonstrates that they come to fill important discourse functions that are more or less used in all languages. These are functions such as narration, providing background information, making predictions, stating intentions and so on. The nature of the content and function of these categories reveals much about human cognition and predilections for interactive communication.

> Without a substantive semantic theory that includes treatments of both open-class and closed-class items and the semantic relations between them (which Bybee and McClelland (2005: 404) admit they lack), answers based on grammaticalization cannot be the end of the story. A substantive semantic theory, at least according to the evidence with which I am familiar (FL: Chapters 9–12), will have to presuppose just the sort of innate grain to human thought and motivation that Bybee and McClelland's radical associationist approach to mind denies. (Jackendoff 2007: 373).

Jackendoff is correct that a good semantic theory is needed as a foundation to explanation of the semantic changes in grammaticalization. This is exactly what grammaticalization research is developing based on the empirical evidence. For the domains of tense, aspect and modality such a theory is well-developed, based on extensive cross-linguistic evidence on language change (Bybee, Perkins and Pagliuca 1994; Traugott and Dasher 2002). The interaction of closed and open class items has been one of the foci of this research, especially in cases where it is relevant, as in verbal aspect. Similar research on determiners (Givón 1981; Heine 1997; Hopper and Martin 1987), spatial relations (Heine, Claudi and Hünnemeyer 1991; Svorou 1993), and other domains provide a rich base for understanding the semantics of natural language and the interaction with pragmatics and discourse structure. Bybee and McClelland (2005: 404) say it is difficult to specify in a connectionist simulation the representation of the meaning of an utterance, not that cognitive-functional approaches to linguistics or grammaticalization have no semantic theory. Finally, in response to the last comment in the passage quoted above, presupposing an "innate grain to human thought and motivation" is not the same as presupposing that there are innate linguistic universals. In any case, we would suppose there is a natural grain to human thought and a natural basis for human motivation and refer the reader to the previous section for our views on their possible innateness.

6. On the Limitations of Connectionist Models

In his reply, Jackendoff makes several statements about the limitations of connectionist models relative to other approaches. First, he says:

> All standard linguistic theories give us a handle on how to analyze sentences like (8), which is (to be just a little cruel) the opening sentence of Bybee and McClelland's article:

(8) There is a range of views about the relationship between language as an abstract system and what people actually say when they talk – what de Saussure called 'langue' and 'parole'.

Despite over twenty years of research on connectionist modeling, no connectionist model comes close. (2007: 373)

While we would not wish to claim connectionist approaches have succeeded fully in addressing the processing of complex sentences, we would also point out that such sentences pose challenges for other approaches, and our first sentence is a case in point. First of all, there is no consensus among linguists on how this sentence should be represented. One debatable point is whether the underlying structure of this sentence provides any indication of the intended relationship between *langue* and *language as an abstract system* on the one hand and *parole* and *what people say when they talk* on the other. In our view, any approach that fails to establish this relationship can't really be said to have represented the underlying structure of the sentence. Second, the determination that establishing such a relationship is the correct thing to do in this case is not simply structure dependent, and any process that is based on abstract rules of the kind found in standard linguistic theories won't in general find the correct parse. Consider a slight variant of the sentence above:

(8′) There is a range of views about the relationship between language as an abstract system and what people actually say when they talk – what Jones calls usage and performance.

In (8′), it appears likely that most readers would think that Jones (a fictive character!) is using two terms to name aspects of what people actually say when they talk, rather than to name language on the one hand and what people actually say on the other. We submit that our sentence worked – even (to be a little cruel ourselves) fooling Jackendoff into thinking it is actually an easy one to parse – because its readers (the readership of TLR) all know a great deal about these issues in general and de Saussure's approach to them in particular. If the reader were, say, an adult American citizen with a BS in biology who never studied linguistics or French (yes, the meanings of the words 'langue' and 'parole' provide important constraints on the interpretation of this sentence), the sentence would at best be ambiguous.

Jackendoff goes on to detail four specific limitations he sees in connectionist/neural network models:

a. A neural network has no way to represent the relationships among items in a sentence being perceived – to build structure on line – because there is no independent working memory. (374)

b. A neural network has no way to represent multiple occurrences of the same item in a sentence. (374)
c. A standard neural network cannot encode a general relation such as *X is identical with Y*, *X rhymes with Y*, or *X is the past tense of Y*. (374)
d. In neural networks, long-term memories are encoded in terms of connection strengths among units in the network, acquired through thousands of steps of training. This gives no account of one-time learning of combinatorial structures, such as the meaning of *I'll meet you for lunch at noon*, a single utterance of which can be sufficient to cause the hearer to show up for lunch. (375)

These points provide a useful starting point for reply, since they indicate several specific issues that any model of language processing (or systematic cognition) should address. They also provide a good starting place for a review of an extensive connectionist literature that has in fact addressed them. Evidently, this is a literature of which Jackendoff (and perhaps others who have rejected connectionist approaches) appears to be completely unaware. Each of these issues has in fact been the subject of intensive research within the connectionist literature. Here we attempt only to point to some of the key papers in the relevant literature. We stress our own work and other efforts closely related to it but in fact there have been many who have addressed these and other related issues (For a synthetic consideration from a perspective similar to ours, see Feldman 2006). In any case, here are brief responses to each of these issues.

a. A neural network has no way to represent the relationships among items in a sentence being perceived – to build structure on line – because there is no independent working memory. (374)

A model by St. John and McClelland (1990) addressed this issue, and subsequent models by many others have advanced the approach. St. John and McClelland proposed a representation called 'the Sentence Gestalt' (SG) which is built up as each word of a sentence is heard and which provides the context in which each successive word is interpreted. After each word the representation is updated, providing a basis for answering questions about the event described in the sentence. Jackendoff would surely object, because the SG uses learned distributed representations that do not make the structure of the sentence explicit. Yet the SG provides a representation that allows the model to make correct answers to questions about the fillers of each of the roles implicated in the sentence. It also resolves ambiguities, selects specific contextually-appropriate meanings for abstract words (e.g., a container filled with water vs. a container filled with apples), fills in missing arguments (e.g., plausible instruments for 'The boy ate the cereal' vs. 'The boy cut the steak') and addresses many other aspects of what we take to be essential for any mechanism claimed to have understood the contents of a simple sentence.

b. A neural network has no way to represent multiple occurrences of the same item in a sentence.

Jackendoff uses the example *My cat can beat up your cat*, in which the two instances of cat have to be distinguished. This issue has been addressed in several connectionist models, particularly those of Miikkulainen and Dyer (1989) in which specific entities were assigned unique identifying tags as well as inheriting the general properties of the class. In the Story Gestalt model of St. John (1992), individual human names could be used in sentences to assign individuals to roles in simple stories, e.g., 'When John entered the restaurant he was greeted by Mark, who escorted him to a table'. Assignment of specific names to participants in the stories used was arbitrary and varied from story to story. Even so, within a story, an initial sentence such as this allowed the model to track John as customer and Mark as a member of the restaurant staff, giving, e.g., plausible answers to questions such as 'who paid the bill' even if the payment of the bill were not explicitly mentioned. (In this example, we, like Miikkulainen and Dyer, are treating 'John' as 'a person named John' and Mark as 'a person named Mark', no different in principle from my cat and your cat).[1]

c. A standard neural network cannot encode a general relation such as *X is identical with Y*, *X rhymes with Y*, or *X is the past tense of Y*.

This item harkens back to issues addressed under productivity and systematicity above, especially when Jackendoff goes on to say "Connectionists (including Bybee and McClelland 2005:403) claim that there are no such general relations – there are only family resemblances among memorized items, to which novel examples are assimilated by analogy." This point reflects a deep misunderstanding of our views, as already discussed extensively, but to recap: Our view is that all degrees of generality/systematicity are encompassed in our philosophy, including the most extreme degrees. Of more relevance here, Jackendoff's point also reflects a deep misunderstanding of the extent to which a connectionist model can learn and indeed can even impose a pressure toward systematicity from exposure to examples of a particular relationship. Here is one case in point. Consider the relations 'identical to' and 'negation of', applied to patterns of 1's and 0's. In their paper introducing the back-propagation learning algorithm, Rumelhart, Hinton and Williams (1986) presented a network that learned this relation from examples. The network was trained with *X-relation-Y* triples, in which *X* was a binary pattern of length *N*, the relation

1. Jackendoff raises the further issue of linking each cat to its respective referent; we do not think this raises any new matters of principle over and above those that we do address, which we take to be the core of the issue in question.

was specified with a single bit with value 0 for 'identical to' and 1 for 'negation of', and *Y* was the correct pattern having the indicated relation to *X*. The network learned from examples to complete probes consisting of *X* plus the relation bit with the appropriate corresponding output. We have conducted an extension of this simulation, using training examples chosen at random from the space of possible input patterns. A variety of values of N were considered, ranging from 1 to 30. The number of examples required to learn the relation was on the order of 1000, independent of the pattern length N. Considering that the number of possible examples is 2^{N+1}, the fraction of possible examples required for learning grows very small very quickly. For the case of N = 30, the fraction is less than 1/(2,000,000). With small N it looks as though memorization of each case is required; with large N it is apparent that a very general relationship has been learned by the network.

d. In neural networks, long-term memories are encoded in terms of connection strengths among units in the network, acquired through thousands of steps of training. This gives no account of one-time learning of combinatorial structures, such as the meaning of *I'll meet you for lunch at noon*, a single utterance of which can be sufficient to cause the hearer to show up for lunch.

In fact, the issue of one-trial learning in connectionist networks has been the focus of an extensive body of research. A theory called the *Complementary Learning Systems* theory was developed by McClelland, McNaughton and O'Reilly (1995) expressly to address this issue, and that work has led to follow-up on a wide range of specific issues in learning and memory. The problem Jackendoff raises is a general one for cognitive science, and McClelland, McNaughton and O'Reilly addressed it in very general terms. They observed among other things that a patient with a bi-lateral lesion to the medial temporal lobes would in fact understand the sentence 'I'll meet you for lunch at noon' but would be completely unable to retain that information after a moment's distraction from rehearsing it. This and a wealth of information on the neural basis of memory in rodents, primates, and humans lead to the view that there is a special fast-learning system in the medial temporal lobes that sits atop the slow-learning systems that lie in the neocortex. Connectionist models of the type to which Jackendoff refers are thought to be of the kind found in the neocortex, but connectionist networks that are parameterized differently have been the subject of intensive investigation and have been offered by many researchers to address the fast-learning system in the hippocampus. Indeed, the theory is an extension of the early domain-general effort of Marr (e.g., 1971) to develop a general theory of learning, with complementary systems in neocortex and what he called *archicortex* (corresponding to the medial temporal lobes). We stress that the issues here are far more general than language pro-

cessing; they reflect very general constraints on learning systems, namely that there is a need to learn slowly to extract structure from experience and also a need to learn quickly without corrupting what has been gradually learned. Both are required for all aspects of cognition, including language.

7. Conclusion

It is evident that communication between representatives of different approaches to the study of language and cognition is far from ideal. We have each been separately struck by the lack of understanding Jackendoff brings to topics in which we ourselves are far more deeply immersed. We are glad to take a part of the responsibility, and we see this reply as an attempt to address the large gulf that exists between his perspective and ours.

We also see this gulf as understandable in view of the history of the field. The Chomskyan revolution of 50 years ago appeared to sweep away any basis for appealing to general purpose mechanisms, and the habits of mind set in place by Chomsky's insistence on abstract rules – and the need he and others felt to capture them using symbolic forms of computation – have been very hard to overcome. Thus far, the predominant move in response to evidence of graded structure and gradual change appears to be to concede that symbolic rule-like mechanisms only apply in certain cases, but to hold the ground for their full applicability in what are viewed as core cases. What we would argue is that the continued insistence on the fundamental correctness of this view blocks access to a wealth of ideas that will ultimately increase our understanding of language, as is it acquired, represented, and used in people's minds and as it changes over time as people speak and listen.

Stanford University
University of New Mexico

References

Albright, Adam, and Bruce Hayes (2003). Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition* 90: 119–161.

Berko, Jean (1958). The child's learning of English morphology. *Word* 14: 150–177.

Bybee, Joan L. (1985). *Morphology: A Study of the Relation between Meaning and Form*. Philadelphia: John Benjamins.

— (1995). Regular morphology and the lexicon. *Language and Cognitive Processes* 10: 425–455.

— (1998). A functionalist approach to grammar and its evolution. *Evolution of Communication* 2: 249–278.

— (2001). *Phonology and Language Use*. Cambridge: Cambridge University Press.

— (2003a). Cognitive processes in grammaticalization. In *The New Psychology of Language, Volume II*, Michael Tomasello (ed.), 145–67. Mahwah: Erlbaum.

— (2003b). Mechanisms of change in grammaticization: The role of frequency. In *Handbook of Historical Linguistics*, Richard D. Janda and Brian D. Joseph (eds.), 602–623. Oxford: Blackwell Publishers.

— (2006). From usage to grammar: The mind's response to repetition. *Language* 82: 529–551.

Bybee, Joan, and Östen Dahl (1989). The creation of tense and aspect systems in the languages of the world. *Studies in Language* 13: 51–103.

Bybee, Joan, and David Eddington (2006). A usage-based approach to Spanish verbs of 'becoming'. *Language* 82: 323–355.

Bybee, Joan, and Paul J. Hopper (eds.) (2001). *Frequency and the Emergence of Linguistic Structure*. Amsterdam: John Benjamins.

Bybee, Joan and James L. McClelland (2005). Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *The Linguistic Review* 22: 381–410.

Bybee, Joan, and Carol Lynn Moder (1983). Morphological classes as natural categories. *Language* 59: 251–270.

Bybee, Joan, and William Pagliuca (1987). The evolution of future meaning. In *Papers from the VIIth International Conference on Historical Linguistics*, A.Giacalone Ramat, O. Carruba, and G. Bernini (eds.), 109–122. Amsterdam: John Benjamins.

Bybee, Joan, William Pagliuca, and Revere Perkins (1991). Back to the future. In *Approaches to Grammaticalization, Volume II*, Elizabeth C. Traugott and Bernd Heine (eds.), 17–58. Amsterdam: John Benjamins.

Bybee, Joan, Revere Perkins, and William Pagliuca (1994). *The Evolution of Grammar: Tense, Aspect and Modality in the Languages of the World*. Chicago: University of Chicago Press.

Bybee, Joan, and Joanne Scheibman (1999). The effect of usage on degrees of constituency: The reduction of *don't* in English. *Linguistics* 37: 575–596.

Bybee, Joan, and Dan I. Slobin (1982). Rules and schemas in the development and use of the English past tense. *Language* 58: 265–289.

Cinque, Guglielmo (1999). *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford: Oxford University Press.

Croft, William (1991). *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information*. Chicago: University of Chicago Press.

— (2001). *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

Elman, Jeffrey L. (2005). Connectionist models of cognitive development: Where next? *Trends in Cognitive Sciences* 9 (3): 111–117.

Feldman, Jerome A. (2006). *From Molecule to Metaphor: A Neural Theory of Language*. Cambridge, MA: MIT Press.

Givón, T. (1981). On the development of the numeral 'one' as an indefinite marker. *Folia Linguistica Historica* 11 (1): 35–53.

Goldin-Meadow, Susan (2005). What language creation in the manual modality tells us about the foundations of language. *The Linguistic Review* 22: 199–225.

Haiman, John (1994). Ritualization and the development of language. In *Perspectives on Grammaticalization*, William Pagliuca (ed.), 3–28. Amsterdam: John Benjamins.

Hay, Jennifer (2001). Lexical frequency in morphology: Is everything relative? *Linguistics* 39: 1041–1070.

— (2002). From speech perception to morphology: Affix ordering revisited. *Language* 78: 527–555.

Heine, Bernd (1997). *Cognitive Foundations of Grammar*. Oxford: Oxford University Press.

Heine, Bernd, Ulrike Claudi, and Friederike Hünnemeyer (1991). *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.

Heine, Bernd, and Mechthild Reh (1984). *Grammaticalization and Reanalysis in African Languages*. Hamburg: Buske.

Hopper, Paul J. (1987). Emergent grammar. In *Proceedings of the 13th Annual Meeting of the Berkeley Linguistic Society*, Jon Aske, Natasha Beery, Laura Michaelis, and Hana Filip (eds.), 139–157. Berkeley: University of California at Berkeley.

Hopper, Paul J., and Janice Martin (1987). Structuralism and diachrony: The development of the indefinite article in English. In *Papers from the VIIth International Conference on Historical Linguistics*, A.Giacalone Ramat, O. Carruba, and G. Bernini (eds.), 295–304. Amsterdam: John Benjamins.

Jackendoff, Ray (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.

— (2007). Linguistics in cognitive science: the state of the art. *The Linguistic Review*, this issue: 347–401

Lupyan, Gary, and James L. McClelland (2003). Did, made, had, said: Capturing quasi-regularity in exceptions. *Proceedings of the Annual Meeting of the Cognitive Science Society* 25: 740–745.

MacWhinney, Brian, and Jared Leinbach (1991). Implementations are not conceptualizations: Revising the Verb Learning Model. *Cognition* 40: 121–157.

Marcus, Gary F. (2001). *The Algebraic Mind: Integrating Connectionism and Cognitive Science*. Cambridge, MA: MIT Press.

Marr, David (1971). Simple memory: A theory for archicortex. *Philosophical Transactions of the Royal Society of London Series B* 262: 23–81.

McClelland, James L., Bruce L. McNaughton, and Randall C. O'Reilly (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review* 102: 419–457.

McClelland, James L., and Karalyn Patterson (2002). Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Sciences* 6 (11): 465–472.

McClelland, James L., David E. Rumelhart, and the PDP Research Group (eds.) (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume II*. Cambridge, MA: MIT Press.

Miikkulainen, Risto, and Michael G. Dyer (1989). A modular neural network architecture for sequential paraphrasing of script-based stories. *International Joint Conference on Neural Networks* 2: 49–56.

Newmeyer, Frederick J. (2005). *Possible and Probable Languages: A Generative Perspective on Linguistic Typology*. Oxford: Oxford University Press.

Pinker, Steven (1991). Rules of language. *Science* 253: 530–555.

Pinker, Steven, and Michael T. Ullman (2002). The past and future of the past tense. *Trends in Cognitive Sciences* 6 (11): 456–463.

Plaut, David C., James L. McClelland, Mark S. Seidenberg, and Karalyn Patterson (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review* 103: 56–115.

Roberts, Ian, and Anna Roussou (2003). *Syntactic Change: A Minimalist Approach to Grammaticalization*. Cambridge: Cambridge University Press.

Romaine, Suzanne (1995). The grammaticalization of irrealis in Tok Pisin. In *Modality in Grammar and Discourse*, Joan Bybee and Suzanne Fleischman (eds.), 389–427. Amsterdam: John Benjamins.

Rumelhart, David E., Geoffrey E. Hinton, and Ronald J. Williams (1986). Learning representations by back-propagating errors. *Nature* 323: 533–536.

Rumelhart, David E., and James L. McClelland (1986). On learning the past tenses of English verbs. In McClelland, Rumelhart et al. (1986), 216–271.

Rumelhart, David E., James L. McClelland, and the PDP Research Group (eds.) (1986). *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Volume I*. Cambridge, MA: MIT Press.

St. John, Mark F. (1992). The Story Gestalt: A model of knowledge-intensive processes in text comprehension. *Cognitive Science* 16 (2): 271–306.

St. John, Mark F., and James L. McClelland (1990). Learning and applying contextual constraints in sentence comprehension. *Artificial Intelligence* 46: 217–257.

Seidenberg, Mark S., and James L. McClelland (1989). A distributed, developmental model of word recognition and naming. *Psychological Review* 96: 523–568.

Skousen, Royal (1989). *Analogical Modeling of Language*. Dordrecht: Kluwer Academic Publishers.

— (1992). *Analogy and Structure*. Dordrecht: Kluwer Academic Publishers.

Slobin, Dan I. (1985). Introduction: Why study acquisition crosslinguistically? In *The Crosslinguistic Study of Language Acquisition. Volume I: The Data*, Dan I. Slobin (ed.), 3–24. Hillsdale, NJ: Lawrence Erlbaum.

— (1997). The origins of grammaticizable notions: Beyond the individual mind. In *The Crosslinguistic Study of Language Acquisition, Vol. 5: Expanding the Contexts*, Dan I. Slobin (ed.), 1–39. Mahwah, NJ: Lawrence Erlbaum.

Svorou, Soteria (1993). *The Grammar of Space*. Amsterdam: John Benjamins.

Talmy, Leonard (1985). Lexicalization patterns: Semantic structure in lexical forms. In *Language Typology and Syntactic Description, Vol. 3: Grammatical Categories and the Lexicon*, Timothy Shopen (ed.), 36–149. Cambridge: Cambridge University Press.

— (1988). The relation of grammar to cognition. In *Topics in Cognitive Linguistics*, Brygida Kudzka-Ostyn (ed.), 165–206. Amsterdam: John Benjamins.

Traugott, Elizabeth C. (1982). From propositional to textual and expressive meanings: Some semantic-pragmatic aspects of grammaticalization. In *Perspectives on Historical Linguistics*, Winfred P. Lehmann and Yakov Malkiel (eds.), 245–271. Amsterdam: John Benjamins.

— (1989). On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language* 65: 31–55.

— (1995). Subjectification in grammaticalization. In *Subjectivity and Subjectivisation*, Dieter Stein and Susan Wright (eds.), 37–54. Cambridge, UK: Cambridge University Press.

Traugott, Elizabeth C., and Richard B. Dasher (2002). *Regularity in Semantic Change*. Cambridge, UK: Cambridge University Press.