SSLA, **24**, 237–247
DOI: 10.1017.S0272263102002085

NEWS FLASH—HUME STILL DEAD

Lynn Eubank
University of North Texas

Kevin R. Gregg
St. Andrew's University

Ellis refers to a wide range of data supporting the uncontroversial claim that there are frequency effects in linguistic and other behavior. He further resurrects the long-discredited claim that language acquisition consists of frequency-based abstraction of regularities from input. What he fails to do is to show how the former claim leads to the latter, or indeed to show any evidence for this claim. Ellis ignores fundamental and well-known problems—the poverty of the stimulus, cases of instantaneous acquisition, evidence for innate knowledge, evidence that even children regularly ignore perceptual input, the compositionality of concepts, and the systematicity of language—that make his frequency-based account of acquisition a nonstarter.

Linguistic behavior, like other behavior, often shows frequency effects. Ellis takes this uncontroversial, indeed banal, fact and tries to use it to explain the nature of grammar, language acquisition, synchronic variation, sociolinguistics, diachronic language change, the representation of concepts, life, the universe, and everything. The good news is that if Ellis is correct, then we have a beautifully general explanation for an impressively wide range of fundamentally important linguistic phenomena. The bad news is that there is not the slightest reason in the world to think that he is correct. We do not have the space to deal with all of Ellis's claims, but we will look at a couple of important areas—language acquisition, linguistic representation, and rules—and it will be seen that his accounts of these areas are risibly inadequate.

Address correspondence to: Lynn Eubank, P.O. Box 520, Angel Fire, NM 87710; e-mail: lynneubank@yahoo.com.

© 2002 Cambridge University Press 0272-2631/02 $9.50

https://doi.org/10.1017/S0272263102002085 Published online by Cambridge University Press

LANGUAGE LEARNING AND THE "POWER LAW"

Ellis makes much of the so-called power law of practice. However, the areas of (language) learning where the power law is relevant are severely limited. First and foremost, it is obvious that children know which form-function pairings are possible in human-language grammars and which are not, regardless of exposure. Examples of this kind abound; to name but one, children know that PERSON—very frequent in input—can be linked to function morphemes but that ROUGH SHAPE—also very frequent in input—cannot be so linked. The child's internal contribution to language acquisition is effectively to restrict the hypothesis space of what comprises a possible human-language grammar.

Additionally, of course, there are the countless cases of instantaneous learning (McCarthy, 1988; Sterelny, 1990; Touretzky & Hinton, 1988), including language learning, such as so-called fast mapping (Bloom, 2000; Carey, 1978). One does not need to pet a thousand cats to discover that cats are furry; just as Carey's subjects, say, did not need a thousand chromium-colored cups to learn (at least part of) what *chromium* means.

The power law also fails to tell us anything about knowledge that comes about in the absence of exposure (i.e., a frequency of zero) including knowledge of what is *not* possible. Infants seem to have experience-independent knowledge of "folk physics," for instance (Baillargeon, 1995; Spelke, 1994); they know (or believe), in the absence of experience, that objects do not just disappear and reappear or move without impulsion. Ellis aptly points to infants' ability to do statistical analyses of syllable frequency (Saffran, Aslin, & Newport, 1996); but of course those infants have not *learned* that ability. What needs to be shown is how infants uniformly manage this task: Why do they focus on syllable frequency (instead of some other information available in exposure)? How do they know what a syllable is in the first place, given cross-linguistic variation? Much the same is true for other areas of linguistic import—for example, the demonstration by Marcus, Vijayan, Bandi Rao, and Vishton (1999) that infants can infer rules. Of course, work by Crain, Gordon, and others (Crain, 1991; Gordon, 1985) showed early grammatical knowledge in cases where input frequency could not possibly be appealed to. Landau and Gleitman (1985) even documented lexical acquisition *in spite of* frequent input, in which a blind child acquired (her own interpretation of) verbs like *look* despite frequent training under a different interpretation. All of which is to say, for starters, that such claims as that "for language learners to be accurate and fluent in their generalizations they need to have processed sufficient exemplars . . . " (p. 167) are either outright false or else true only vacuously (if "sufficient" is taken to range from as low a figure as 1).

Even where one can plausibly appeal to frequency, moreover, the explanatory power of the power law is limited. Larsen-Freeman (1976), for instance, may have successfully explained the acquisition order of English grammatical morphemes by appealing to their relative frequency in the input, but this is a

https://doi.org/10.1017/S0272263102002085 Published online by Cambridge University Press

far cry from explaining the acquisition *tout court* of those morphemes. This lack of explanatory power is a problem for Ellis's favored accounts of language learning, frequency-based accounts like the Competition Model (e.g., Bates & MacWhinney, 1989), and connectionist models.

The Competition Model, for instance, relies heavily on the frequency (and reliability) of so-called cues. The problem is that it is nowhere explained just what a cue is or what could be a cue; which is to say that the concept is totally vacuous (Gibson, 1992). In the absence of any principled characterization of the class of possible cues, an explanation of acquisition that appeals to cue frequency is doomed to arbitrariness and circularity. (The same goes, of course, for such claims as Ellis's [p. 175] that "the real stuff of language acquisition is the slow acquisition of form-function mappings," in the absence of any criterion for what counts as a possible function and what counts as a possible form.)

Ellis's appeal to connectionist models of language learning is equally shaky. To start with, it is not yet at all clear whether such models are in any way relevant to human learning. Back-propagation, for instance, which is the most widely used learning algorithm in connectionist models (including that of Ellis & Schmidt, 1997, 1998), is widely recognized to have no analogue in human learning and indeed to be physiologically implausible (Clark, 1993; Elman et al., 1996; Smolensky, 1988; Stich, 1994). Back-propagation, where connection strengths are altered on the basis of computations of the deviation from the target value—a kind of negative evidence—was introduced as a mathematical solution to an engineering problem, not as a theoretically motivated solution to a problem in human learning (Smolensky). This is to say that any putative success of a connectionist model can be generalized to human learning only on the basis of wholly unsupported assumptions about human learning.

What sort of successes have connectionist models shown so far? *Pace* Ellis, these models have not shown an ability to generalize beyond their training sets, and they do not produce new categories or modules (Cummins, 1991; Marcus, 1998a). Thus, for example, the plural-learning model of Ellis and Schmidt (1997, 1998)—again, *pace* Ellis—was never subjected to a wug test and would not have known what to do with a truly novel noun if it had been so tested. The model also did not learn the regular plural prefix because it did not learn any prefixes at all. Rather, it merely learned to associate certain units (stipulated to be nouns) with a certain other unit (stipulated to be the regular plural prefix); but this is a far cry from learning the category prefix. In fact, as far as the model is concerned, the output units could be suffixes, infixes, a few of each, zero conversion, or none of the above. Associations in and of themselves, after all, are not structured or even linearly arranged (let alone hierarchically arranged, as is the case with representations in natural languages). Ellis is thus quite misleading when he tells us (p. 165) that for connectionist networks, "learning the grammatical categories . . . reduces to the analysis of the sequence in which words work in chunks," because the models he discusses do not learn the grammatical categories in the first place. At

https://doi.org/10.1017/S0272263102002085 Published online by Cambridge University Press

best, they learn—to some extent—the positional frequencies of words in strings; calling the clusterings produced by these models "categories" is hopelessly question-begging. As Fodor points out, "the problem of language acquisition is that of how a child learns grammatical structure, not how he learns *correlates* of grammatical structure" (Fodor, 1998b, p. 150, emphasis in original; cf. Marcus, 1998b).

Ironically, connectionist models are, in fact, both absurdly innatist and absurdly modular (Chomsky, 1998; Marcus, 1998b). That is, a given network's input and output units are determined *deus ex machina* by the experimenter (i.e., they are innately given), and they are totally incapable of dealing with data outside the training set. This makes them a dangerously weak evidential reed for Ellis's associationism to lean on. (For more detailed discussion of connectionist models in SLA, see Gregg, 2001.)

In short, the associationist account Ellis proposes as the overall explanation of how language is acquired (a) fails to provide even an unsatisfactory account of most aspects of language competence, (b) fails to provide a realistic model for the acquisition of those aspects of language competence for which a satisfactory associationist account might at least be plausible, and (c) fails totally to explain the process of learning (see below).

LINGUISTIC COMPETENCE AND LINGUISTIC PERFORMANCE (OH, GOD, YET AGAIN!)

It really should go without saying by now that linguistic theory makes a distinction between knowledge and the behavior based on that knowledge, between competence and performance. This distinction has often been rejected (without plausible grounds) and more often ignored or glossed over, but no coherent linguistic theory has managed to do without it. This is completely unsurprising, because it is merely the application to the domain of language of the fundamental distinction made by all sciences between underlying phenomena and observable data (Gregg, 1993). Ellis seems to have trouble with this distinction. He confuses grammaticality, for instance, with grammaticality judgments (p. 160–163)—and both, apparently, with acceptability judgments. Again, he argues that "The knowledge of a speaker-hearer cannot be understood as a fixed grammar but rather as a statistical ensemble of language experiences" (p. 162). This characterization of linguistic competence is bizarre, to say the least; but more to the point, it is unsupported by anything Ellis says or alludes to. (It also contradicts Ellis's favored grammatical theory, Construction Grammar [cf. Goldberg, 1999, p. 200].) As Chomsky pointed out 45 years ago, in a passage Ellis himself cites, there is no known connection between input frequency and grammatical structure. Any native speaker of English knows, for example, that a sentence of the form *He may have been being followed* is possible whereas *The students want me to help each other* is not, but

https://doi.org/10.1017/S0272263102002085 Published online by Cambridge University Press

the frequencies with which a speaker hears the two sentence types are identical, that is, zero.

For that matter, frequency does not have that much to do with performance, either. Ellis actually says that "language users tend to produce the most probable utterance for a given meaning on the basis of frequencies of utterance representations" (p. 145), but that is just too ludicrous for words. It would mean, so far as we can make any sense of the claim, that a speaker wishing to say that sheepdip makes an excellent dessert topping would consult the relative frequencies in her mental lexicon—or else the frequencies across the population as somehow represented in her mind—to establish what the best choices of words were for *sheepdip*, *excellent*, and so on. (Note in this regard that reliance on semantic feature combinations, as in Ellis's discussion of "colorless green ideas . . . ," entirely misses the point that the features were supplied, again *deus ex machina*, by the researchers.)

Ellis's point is well taken that a hearer makes use of relative frequencies to disambiguate an input utterance. However, note (a) that we can and do override frequency-based interpretations of ambiguous input (*I took the plane to Chicago, where it now stands in my garden*), and these decisions are based on inferential considerations that have little or nothing to do with frequency; and (b) that the choice we make among interpretations is constrained by the grammar, and these constraints are not statistical. The sentence *The border guard asked the traveler for a picture of himself*, for instance, would no doubt be interpreted by most of us as meaning that the guard asked for a picture of the traveler. It might even be true that we make this interpretation on the basis of input frequency. (Actually, we do not believe that for a minute: Frequency of what? *Ask-for* constructions that imply this relationship? Sentences with *himself* that imply this relationship? Instances we have experienced of guards asking travelers for pictures?) In any case, though, we can make this interpretation because the grammar of English sanctions it. The Japanese equivalent to this sentence, on the other hand, cannot mean the guard asked for the traveler's picture, and so a Japanese hearer does not have the freedom to make that interpretation. It is not that the sentence tends not to have that meaning; it cannot have it, *punkt*.

In short, Ellis has nothing to say about central areas of linguistic competence and little to say about the rest.

CONCEPTS AND PROTOTYPES, RULES AND TENDENCIES

For Ellis, rules are really just extremely high correlations. As Ellis and Schmidt (1998) put it, "From an associative perspective, regularity and frequency are essentially the same factor under different names" (p. 329). A rule-based explanation, for instance, of the fact that the past tense of *exacerbate* is *exacerbated* would be that natives have, on the basis of limited input, induced a rule,

https://doi.org/10.1017/S0272263102002085 Published online by Cambridge University Press

which then becomes the deductive basis for inferring the past tense of newly encountered forms. An Ellis-style associationist would say that it is *exacerbated* not because we have a past-tense rule but because the past tense of overwhelmingly many other verbs tends to be *-ed*. There are ways of choosing between these two accounts (see Pinker, 1999).

For one thing, rules support counterfactuals; generalizations do not. Thus, given a past-tense rule, we can claim that if *zoop* were an English verb, then it would have a past-tense form; further, we know that if *zoop* has no past-tense form, it is not a verb. Of course, in the case of English past tense or plural, the input data are so voluminous that an associationist could plausibly claim the inference to *zooped* as an inductive generalization rather than a deductive inference. This option becomes less available, however, in other cases: (a) early acquisition, for example, child L1 acquisition, where irregulars are type- and token-frequent and yet a rule seems to develop quickly (Berko, 1958); (b) cases where the most frequent form is not the default (i.e., regular) form (Clahsen, Rothweiler, Woest, & Marcus, 1992; but cf. Hahn & Nakisa, 2000); and (c) correct performance with rarely or never used grammatical structures, for example, "he may have been being followed" (Wexler, 1991), compared to quite frequent errors on vastly more frequent structures, such as minimal distance effects on number agreement.

Of course it is precisely because rules have a deductive structure that one can have instantaneous learning without the trial and error involved in connectionist learning. With the English past-tense rule, one can instantly determine the past-tense form of *zoop* without any prior experience of that verb, let alone of *zooped* (unlike, say, the model of Ellis & Schmidt, 1998, which could only approach the supposedly correct plural form for the test item and only after repeated exposures to the singular form followed by repeated exposures to the plural form, along with back-propagated comparisons). If all we know is that John zoops wugs, then we know instantaneously that John zoops, that he might have zooped yesterday and may zoop tomorrow, that he is a wug-zooper who engages in wug-zooping, that whereas John zoops, two wug-zoopers zoop, that if he is a Canadian wug-zooper he is either a Canadian or a zooper of Canadian wugs (or both), and so forth. We know all this without learning it, without even knowing what *wug* and *zoop* mean. A frequency-equals-regularity account would need to appeal to a whole congeries of associations, between a large number of pairs like *rum-runner* ~ *runs rum*, *piano-tuner* ~ *tunes pianos*, and so forth, but not like *harbor-master* ~ *masters harbors*, *pot-boiler* ~ *boils pots*, *kingfisher* ~ *fishes kings*, and so forth, or a roughly equal number of pairs like *purple people-eater* meaning purple people or purple eater.

This lack of deductive structure in Ellis's associationist mind raises problems for his use of a prototype account of concepts, which in turn raises problems for his idea of language and language acquisition. Just as he fails to distinguish grammaticality judgments from grammaticality, so he fails to distinguish prototype effects from prototypes and categorization from categories. The confusions are serious.

https://doi.org/10.1017/S0272263102002085 Published online by Cambridge University Press

According to Ellis (p. 146), "We know that natural categories are fuzzy rather than nomothetic." In fact, we "know" no such thing; the nature of categories is the subject of hot debate in philosophy and cognitive science, and the putative fuzziness of categories is by no means an established fact. Similarity is not a matter of mere feature counting either, assuming that one knew what features were; in Ellis's hands the term is vacuous. (Of course, if similarity really were a matter of feature counting, there would be no fuzziness, as the number of features is not fuzzy, at least if one knew what features were.) Even young children, for instance, show so-called essentialist (as opposed to the associationist's perceptual) thinking: Children are perfectly willing to say that a coffee pot that has been refabricated to look like a birdfeeder is now a birdfeeder (Pinker, 1997) but will stubbornly deny that a skunk altered by scientists to look and sound exactly like a beaver is now a beaver or that a dog with its insides removed is a dog (Gelman & Wellman, 1991; Keil, 1989). This suggests that even young children have ideas as to the essential nature of some categories, features be damned. Gopnik and Sobel (2000) have shown that even by 3 years of age, "children will override perceptual information in favor of causal information when they classify objects" (Gopnik, 2000, p. 304).

Ellis seems to be arguing that because there are prototype effects—so that we all tend to accept a robin as birdier than a penguin—concepts must therefore be fuzzy. This does not follow, though. The fact that (at least sometimes) we find it hard to decide in which category to put an individual (whales, epicene youths, olives) says nothing about whether or not the categories are fuzzy; whales are mammals, penguins are birds, olives are fruits, *basta*. Indeed, the very ubiquity of prototype effects suggests that there is something unsatisfactory about the idea of prototypes. After all, 3 is reliably taken to be a more prototypical odd number than 27 (Armstrong, Gleitman, & Gleitman, 1983), yet one would be hard put to say what "features" 3 has that 27 lacks, or what is fuzzy about the concept ODD NUMBER.

The biggest problem with prototypes, however, is that prototype accounts fail to account for compositionality (Fodor 1998a, 2001). If, for instance, I know that X is a pet fish, I automatically know, deductively, that X is a pet and that X is a fish (and if I know that X is a pet and that X is a fish, I automatically know that it is a pet fish). In other words, the concept PET FISH is composed from the concepts PET and FISH; it is the intersection of the two extensions. If I know that a goldfish is a prototypical pet fish, though, it does not follow—indeed, it is not correct—that a goldfish is a prototypical pet or that it is a prototypical fish. What this means is that a prototype account of concepts must treat the three concepts as mutually independent primitives, failing to capture the hierarchical relation holding between them. Thus, for instance, Allen and Seidenberg's (1999) model of grammaticality judgments (not grammaticality, *pace* Ellis) represents the word *mother* by using, *inter alia*, the putative features "female-parent," "female," and "parent." In other words, the model—unlike humans—could not deduce that a female parent is a female and a parent, and it could conceive of a female parent that was neither a female nor a parent!

https://doi.org/10.1017/S0272263102002085 Published online by Cambridge University Press

What goes for concepts goes, a fortiori, for linguistic concepts, like NOUN and TRANSITIVE VERB. It follows deductively that if X is a noun, then for example, X has no tense, or that if X is an NP sister of V, then it is part of the VP of which V is head, and so on. It makes no difference whether X is a prototypical noun like *tree* or a nonprototypical noun like *equilibrium*. We have no idea what a prototypical compound noun is or what a prototypical compound is; no doubt it varies by language. However, we do know (in the absence of any experience with at least 99% of the world's languages) that if X is a compound and X is a noun, then X is a compound noun, and that if X is a compound noun, then it is a noun and a compound. Despite a tendency in certain Rossian circles to talk of "squishes," in fact grammatical categories are categorical (Newmeyer, 1998), and they enter into systematic, lawlike relations. As has been often noted and widely recognized (Clark, 1993; Fodor & Pylyshyn, 1988; Hadley, 1994), associationist models of learning are incapable of capturing systematicity. If a language allows a sentence like *John loves Mary*, it also allows *Mary loves John*. Similarly, if one can think that John loves Mary one can also think that Mary loves John; if one can think that *p*&*q*&*r*, one can think that *p*&*q*, and so on. This is true for all human languages and for all human speakers. A linguistic theory (similarly, a theory of mental representation) that posits syntactic rules and principles can explain this universal truth very simply, by appealing to nomological necessity: This kind of systematicity simply follows from general syntactic principles. On an associationist account, however, this universal systematicity is contingent; it literally just so happens that all languages are systematic (Fodor & Pylyshyn), which is to say that Ellis's associationist vision of language is simply not worth taking seriously; not for a moment.

TESTING THEORIES

One of the (nonempirical) virtues of a theory is simplicity; all in all, we would rather have one single cause for a phenomenon than two, two rather than three. Ellis might ask us: Given that frequency clearly plays a role in language acquisition and use, why not grant it *the* role? Well, we would if we could, but we cannot; as we said, simplicity is a nonempirical virtue, and our first duty is to the phenomena. As Crick (cited in Marcus, 1998b) says, "While Occam's Razor is a useful tool in physics, it can be a very dangerous implement in biology" (p. 247). The successes of the germ theory of disease in the nineteenth century led some to propose that all diseases are caused by germs—a beautifully simple theory, whose only defect is that it is wrong; *mutatis mutandis* for Ellis's associationism.

What we want from a theory, first and foremost, is explanatory power. Standard (generative) accounts of language structure and language acquisition—accounts that treat input and input frequency as secondary causal powers at

https://doi.org/10.1017/S0272263102002085 Published online by Cambridge University Press

best—have achieved explanatory coverage of vast amounts of data from countless different languages. This coverage has been attained using a small set of concepts (e.g., c-command) to explain seemingly unrelated phenomena (e.g., coreference, gender marking) in seemingly unrelated languages. In comparison, associationist accounts have come up with next to nothing, as at least some connectionists acknowledge (Smolensky, 1999). Of course, some connectionists (Redington & Chater, 1998; Smolensky, 1988; Touretzky & Hinton, 1988) are perfectly willing, unlike Ellis, to grant an important role to innate, structured knowledge, of the sort posited by standard linguistic theories.

Does this mean that generativist theories are correct? Of course not. The explanatory successes of linguistic theory are numerous only in comparison with those of its competition; but linguists and other cognitive scientists have only scratched the surface of the phenomena. For that very reason we will not emulate Ellis and make sweeping, dogmatic claims about the nature of language to oppose his sweeping, dogmatic claims, some of which we have quoted here (and none of which he supports).

Redington and Chater (1998) argued that it is important to take a hypothesis and push it to the extreme, precisely to test its limits. We agree 100%, which is why we welcome the use of connectionist models, say, in language acquisition research. It would be of real value (perhaps especially to those of us in SLA) to know the extent and limits of connectionist-style learning mechanisms. Pushing a hypothesis in the laboratory, though, is not the same thing as pushing it in the academic marketplace. As Glymour (2000) put it, "Although among animals and very young humans, causal relations must somehow be discovered from associations, associations are not causes. Nothing good can come of confusing the two ideas. Nothing good has" (p. 194).

REFERENCES

Allen, J., & Seidenberg, M. S. (1999). The emergence of grammaticality in connectionist networks. In B. MacWhinney (Ed.), *The emergence of language* (pp. 115–152). Mahwah, NJ: Erlbaum.

Armstrong, S., Gleitman, L., & Gleitman, H. (1983). What some concepts might not be. *Cognition*, *13*, 293–308.

Baillargeon, R. (1995). Physical reasoning in infancy. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 181–204). Cambridge, MA: MIT Press.

Bates, E., & MacWhinney, B. (1989). Functionalism and the Competition Model. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing* (pp. 3–73). New York: Cambridge University Press.

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

Bloom, P. (2000). *How children learn the meanings of words*. Cambridge, MA: MIT Press.

Carey, S. (1978). The child as word-learner. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264–293). Cambridge, MA: MIT Press.

Chomsky, N. (1998, November). *New horizons in the study of language and mind*. Lecture given at the University of Maryland, College Park. (Graduate School Distinguished Lecture Series).

Clahsen, H., Rothweiler, M., Woest, A., & Marcus, G. (1992). Regular and irregular inflection in the acquisition of German noun plurals. *Cognition*, *45*, 225–255.

Clark, A. (1993). *Associative engines*. Cambridge, MA: MIT Press.

https://doi.org/10.1017/S0272263102002085 Published online by Cambridge University Press

Crain, S. (1991). Language acquisition in the absence of experience. *Behavioral and Brain Sciences*, *14*, 597–650.
Cummins, R. (1991). The role of representation in connectionist explanations of cognitive capacities. In W. Ramsey, S. P. Stich, & W. E. Rumelhart (Eds.), *Philosophy and connectionist theory* (pp. 91–114). Mahwah, NJ: Erlbaum.
Ellis, N. C., & Schmidt, R. (1997). Morphology and longer distance dependencies: Laboratory research illuminating the A in SLA. *Studies in Second Language Acquisition*, *19*, 145–171.
Ellis, N. C., & Schmidt, R. (1998). Rules or associations in the acquisition of morphology? The frequency by regularity interaction in human and PDP learning of morphosyntax. *Language and Cognitive Processes*, *13*, 307–336.
Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: MIT Press.
Fodor, J. A. (1998a). *Concepts: Where cognitive science went wrong*. Oxford: Clarendon Press.
Fodor, J. A. (1998b). *In critical condition: Polemical essays on cognitive science and the philosophy of mind*. Cambridge, MA: MIT Press.
Fodor, J. A. (2001). Language, thought, and compositionality. *Mind and Language*, *16*, 1–15.
Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*, 3–71.
Gelman, S. A., & Wellman, H. M. (1991). Insides and essences: Early understandings of the non-obvious. *Cognition*, *38*, 213–244.
Gibson, E. (1992). On the adequacy of the Competition Model. *Language*, *68*, 812–830.
Glymour, C. (2000). Bayes nets as psychological models. In F. C. Keil & R. A. Wilson (Eds.), *Explanation and cognition* (pp. 169–197). Cambridge, MA: MIT Press.
Goldberg, A. E. (1999). The emergence of the semantics of argument structure constructions. In B. MacWhinney (Ed.), *The emergence of language* (pp. 197–212). Mahwah, NJ: Erlbaum.
Gopnik, A. (2000). Explanation as orgasm and the drive for causal knowledge: The function, evolution, and phenomenology of the theory formation system. In F. C. Keil & R. A. Wilson (Eds.), *Explanation and cognition* (pp. 299–323). Cambridge, MA: MIT Press.
Gopnik, A., & Sobel, D. M. (2000). Detecting blickets: How young children use information about novel causal powers in categorization and induction. *Child Development*, *71*, 1205–1222.
Gordon, P. (1985). Level-ordering in lexical development. *Cognition*, *21*, 73–93.
Gregg, K. R. (1993). Taking explanation seriously; or, Let a couple of flowers bloom. *Applied Linguistics*, *14*, 276–294.
Gregg, K. R. (2001, October). *The state of emergentism in SLA*. Plenary address given at the Pacific Rim Second Language Research Forum, Honolulu, HI.
Hadley, R. F. (1994). Systematicity in connectionist language learning. *Mind and Language*, *9*, 247–272.
Hahn, U., & Nakisa, R. C. (2000). German inflection: Single route or dual route? *Cognitive Psychology*, *41*, 313–360.
Keil, F. C. (1989). *Concepts, kinds, and cognitive development*. Cambridge, MA: MIT Press.
Landau, B., & Gleitman, L. R. (1985). *Language and experience*. Cambridge, MA: Harvard University Press.
Larsen-Freeman, D. (1976). An explanation for the morpheme acquisition order of second language learners. *Language Learning*, *25*, 125–135.
Marcus, G. F. (1998a). Can connectionism save constructivism? *Cognition*, *66*, 153–182.
Marcus, G. F. (1998b). Rethinking eliminative connectionism. *Cognitive Psychology*, *37*, 243–282.
Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, *283*, 77–80.
McCarthy, J. (1988). Epistemological challenges for connectionism. *Behavioral and Brain Sciences*, *11*, 44.
Newmeyer, F. J. (1998). *Language form and language function*. Cambridge, MA: MIT Press.
Pinker, S. (1997). *How the mind works*. New York: Norton.
Pinker, S. (1999). *Words and rules*. New York: Basic Books.
Redington, M., & Chater, N. (1998). Connectionist and statistical approaches to language acquisition: A distributional perspective. *Language and Cognitive Processes*, *13*, 129–191.
Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.
Smolensky, P. (1988). The proper treatment of connectionism. *Behavioral and Brain Sciences*, *11*, 1–74.

https://doi.org/10.1017/S0272263102002085 Published online by Cambridge University Press

Smolensky, P. (1999). Grammar-based connectionist approaches to language. *Cognitive Science*, *23*, 589–613.
Spelke, E. (1994). Initial knowledge: Six suggestions. *Cognition*, *50*, 443–447.
Sterelny, K. (1990). *The representational theory of mind: An introduction*. Oxford: Blackwell.
Stich, S. P. (1994). The virtues, challenges, and implications of connectionism. *British Journal for the Philosophy of Science*, *45*, 1047–1058.
Touretzky, D. S., & Hinton, G. E. (1988). A distributed connectionist production system. *Cognitive Science*, *12*, 423–466.
Wexler, K. (1991). On the argument from the poverty of the stimulus. In A. Kasher (Ed.), *The Chomskyan turn* (pp. 252–270). Oxford: Blackwell.

https://doi.org/10.1017/S0272263102002085 Published online by Cambridge University Press