FIVE DECADES AFTER CHOMSKY: AN EXPERIENCED-BASED AWAKENING

Creating Language: Integrating Evolution, Acquisition, and Processing

By Morten H. Christiansen and Nick Chater. Cambridge, MA: MIT Press, 2016. 330 pp. Hardcover, $40.

More than 50 years ago, Noam Chomsky (1959, 1965) threw down the nativist gauntlet on language, and he has withstood potshots from many quarters. His debut involved a critique of B. F. Skinner's (1957) account of language acquisition. Central to Skinner's account was the important influence of the linguistic environment. Fundamental reinforcement principles would effect the association of spoken words with meaning. Quine (1960, 1990/1992) had not yet popularized the Gavagai challenge, in which a child has to determine which of many possible meanings correspond to an uttered word. Similarly, Rescorla and Wagner had not yet broadened the Rescorla–Wagner model of association learning with the concept of reinforcement as information or surprise value (Rizley & Rescorla, 1972). Skinner viewed the child as being reinforced when she realizes that a particular word has a particular meaning. The child learns to distinguish the words *milk* and *spinach* because she is reinforced to reply appropriately to her father's request because she likes one and not the other.

Chomsky's Shadow Sets the Stage

According to Chomsky's critique, children have at their beck and call an unlimited set of sentences, and a child could never acquire this skill based on a paucity of language input along with reinforcement principles. What was needed was an innate Universal Grammar to allow language development in children and language processing in adults. Even though the assumed properties of this facility have changed over the decades, the bottom line is that language could never be learned without being bootstrapped by this innate ability.

With hindsight bias and the intervening five decades of research, we might counter Chomsky's critique with the apparent impact granted by experience and the natural ability to generalize from one event to another. But somehow it has been difficult for the study of language use to distance itself from Chomsky's shadow. Only recently, with increasingly sophisticated empirical and theoretical research and now with the advent of Morten Christiansen and Nick Chater's compelling volume, is there the impression that the field has sunny days ahead. This book, grounded in research on brain, evolution, culture, and language acquisition and use, evolved over a two-decade collaboration. Chapters 2–5 and 7 of the book rework and consolidate several of their previous publications. Two of these publications were *Behavioral and Brain Sciences* articles with peer commentary, which we can expect to have already broadened an existing large scope of study.

Christiansen and Chater (hereafter C&C) maintain a strong case for the negative impact that Chomsky's so-called hidden agenda of generative grammar had on the study of language. From its origins Chomsky's influence has tended to isolate language studies from processing, acquisition, and evolution. As is well known and rehearsed, advocates of generative grammar believe that language performance cannot inform the idealized linguistic competence of a language user. In retrospect, this seems especially odd because generative grammar stimulated many early psycholinguistic experiments looking for its psychological reality. The plethora of click studies alone (in which participants are to locate where a click occurred in a sentence) generated considerable research attempting to uncover the constituent structure of sentences (Freund, 1975). Chomsky saw little value in studying language acquisition because he claimed that language arrived almost full-blown for a 1- or 2-year-old growing up in a typical language environment. Similarly, for Chomsky, there was little of interest in the evolution of language because it came about whole, without an important precedent.

The major contribution of C&C's book is the articulation of the interplay of evolution, processing, and acquisition in a coherent account of language. C&C succeed at their ambitious goal of integrating these three areas of inquiry by describing their interplay and showing how tightly they interact. Figure 1 (Figure 1.6 from their book) illustrates this interaction between these important dimensions for understanding language and how it is acquired and used. Acquisition constrains what can evolve and fits what is learned to the processing mechanism, and processing limits what can be acquired and constrains what can evolve. Evolution fits language to the processing mechanisms and to the learner. I would like to believe that this book and the momentum of its approach should have a large impact and eventually supersede Chomsky's and improve the science of language.

Beyond Universal Grammar

To their credit, C&C take universal grammar (UG) and its falsification very seriously, C&C certainly have not yet converted and probably will not convert many additional followers of the "language is special" camp, as witnessed by some of the commentaries on their two *Behavioral and Brain Sciences* target articles (C&C, 2008). However, their systematic dissection of how UG might have evolved is instructive and worth summarizing here.

They consider two explanations of how the arbitrary principles of UG could be genetically specified: adaptationist and nonadaptationist. The first assumes that UG evolved gradually through natural selection. The second gives a minor role to natural selection and allows the possibility of other events such as a "lucky" mutation or two during evolutionary history. C&C take seriously Pinker and Bloom's adaptationist view that the evolution of the language faculty consisted of many arbitrary constraints to maintain a standardized communication code. C&C describe various limitations of this explanation, one involving the likely occurrence of language change during its evolution. We see language change every day with the arrival of new vocabulary and the disappearance of previously frequent words, such as their examples of the disappearance of *fax* and the recent arrival of *selfie*. We could imagine a similar change in the early appearance of language as people moved to different niches and found new points of discourse. C&C bolster their argument with computer simulations, showing that a fixed language will lead to fixed genes

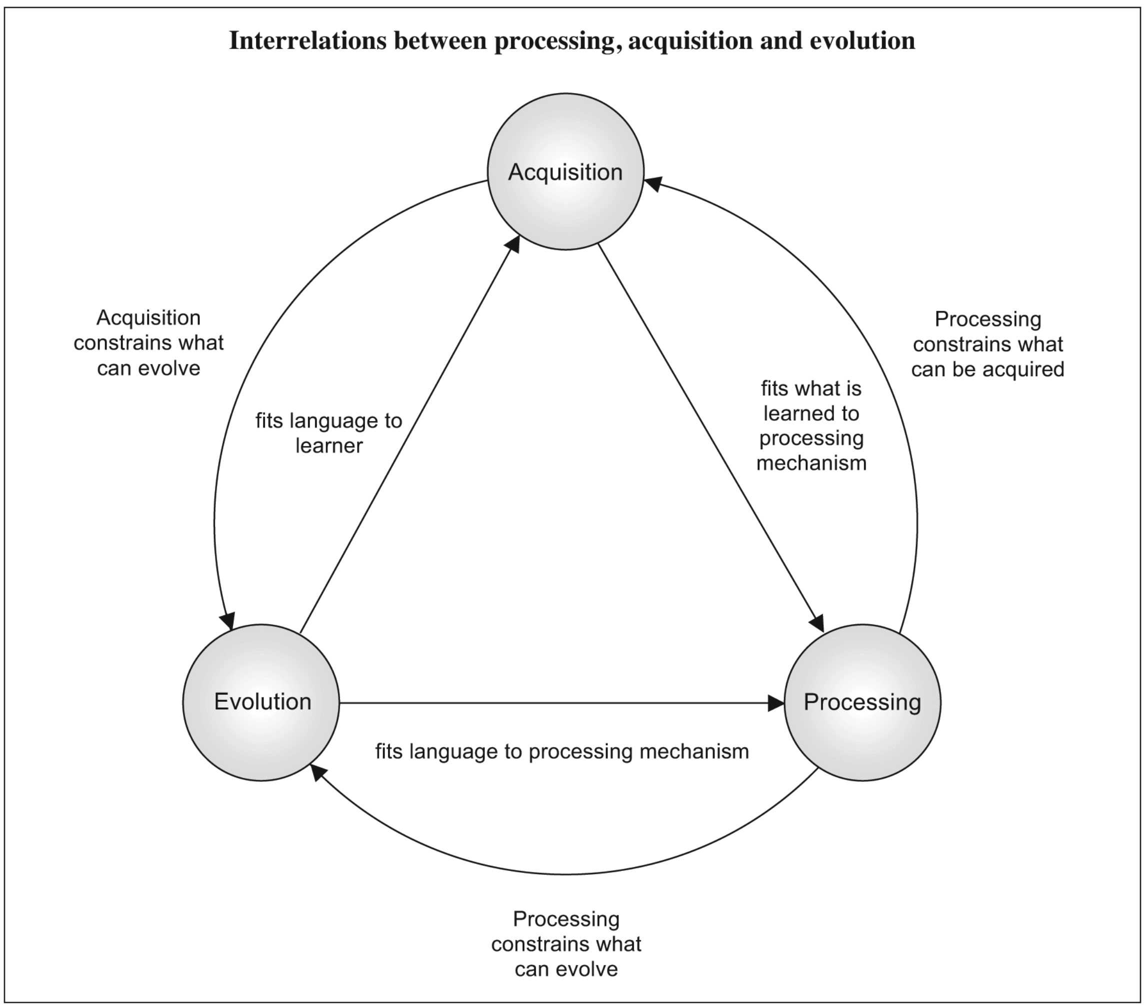

FIGURE 1. The interrelations between the evolution acquisition, and processing of language. From *Creating Language: Integrating Evolution, Acquisitions, and Processing*, by Morten H. Christiansen and Nick Chater, published by The MIT Press, ©MIT 2016 (used with permission)

that will optimize processing of that language, but with a changing language, neutral genes are favored. Neutral genes prepare the language user with multipurpose strategies to handle a language that is continuously changing. (I like the idea of neutral genes because it would be compatible with the possibility of naturally acquired written literacy, which might be conceptualized as just another language change that children from birth onward might master when embedded in the imminent technology and artificial intelligence; Massaro, 2012a.)

C&C's primary critique of the nonadaptationist account is simply that the likelihood of stumbling on a UG fix for language is extremely small. From a somewhat biased perspective, the nonadaptationist account is much too much like "Then a miracle occurs." If early language users were succeeding in language use, why would they need some additional boost from UG to make it possible? Thomas Kuhn's (1962) insights about scientific progress might be helpful here. UG has been consistently modified (patched up?) since almost the time of its conception but will probably not be abandoned until a new paradigm replaces it. Now armed with sophisticated experimental techniques across the life span and the increasing availability of large language corpora, we are witnessing a revolution in the science of what it takes to participate in a linguistic community. Almost daily we are told again that babies are expert pattern recognizers, association engines, and statistical learning machines, important processes for acquiring language (e.g., Wang & Saffran, 2014). C&C add to this dialog by bringing to bear a strong case for experience-based processing and learning.

C&C request that we replace the question of the evolution of language users with the question of the evolution of languages. Using the metaphor of biological evolution, the evolution of languages could follow an analogous path. Given the obvious advantage language would ensure, chatty people would be selected over those reticent to participate in the language game. In addition, C&C offer a huge counterpoint to modularity of the brain. Learning and using language might simply involve exercising existing brain mechanisms in this new domain of gossiping. They discuss important research by Anderson (2010), who analyzed a plethora of functional magnetic resonance imaging experiments to determine which regions of the brain participate in various behaviors. Supporting the claim of nonspecialized processes, the regions involved in language processing are also active in a variety of other nonlinguistic task domains, including attention, memory, reasoning, and action execution.

C&C have argued that our biological adaptation for language has been negligible. The wide diversity of the thousands of languages supports the idea of language adapted to the user rather than biological adaptation accommodating a specific type of language. If this were the case, then it would again make the possibility of naturally acquired literacy using technology more of a possibility because there were not selective adaptations of humans for speech or gesture, but rather speech and gesture had to be accommodated for the language user. Similarly, using technology, written language could be adapted to the infant, toddler, and preschooler (Massaro, 2012a, 2015).

C&C advocate that language learning consists of learning a systematic body of linguistic entities rather than learning specific items in a piecemeal fashion. The interconnections between linguistic items is most apparent in word learning, such as the past tense of verbs. We more easily learn repeated patterns such as *leap* and *leapt* or *sleep* and *slept* than other one-of-a-kind verbs such as *go* and *went* or *lie* and *lay*.

Recursion and Beyond

Before closing their magnum opus, C&C delve into recursion, which appears to be the last bastion of the nativist claim that language is special. All languages putatively have it; this is what is unique about language. Admittedly, there is an unending controversy over what recursion actually is and whether it is truly universal across languages (e.g., Everett, 2005). The authors point out that much of recursion, such as right branching sentences, can be accounted for by a simpler iterative processing. A right branching sentence, "This is the rat that ate the malt that lay in the house that Jack built," is not necessarily recursive. Repeating the construction in this sentence could more simply be generated by iterative processing in terms of a loop that repeats a given structure. A recursive structure must contain self-reference or call itself.

Doubly-embedded recursive sentences such as "The cat the dog the mouse bit chased ran away," do not reduce to simple iteration. And many psycholinguistic experiments have found that this type of sentence is extremely difficult to process and understand. C&C's analysis highlighted for me the irony that the uniqueness claim about recursion in language rests on exactly the recursive sentences that trip up the typical language user. This is similar to the paleontologist telling us that teeth are a special adaptation for food because they crack when they are used to crush frozen food. C&C devote this chapter to interpreting experimental differences in processing different types of sentences as best accounted for by the language users' experience in processing similar structures. The success of this analysis is a major victory for task-specific learning, which is a general principle of learning and not one limited to language processing.

Twenty-First-Century Psycholinguistics

C&C thoroughly review a wide berth of psycholinguistic results, involving corpus analyses and experimental manipulation of linguistic tokens. They offer reasonable explanations of a variety of findings, based primarily on language experience. The authors propose that frequency of exposure is an important force in learning language and ease of language processing. Consider the differences in the difficulty of processing two sentences:

> The reporter that attacked the senator admitted the error.
>
> The reporter that the senator attacked admitted the error.

Experimental research has shown that the first sentence is easier to process than the second, and of course there are different explanations of this finding from both the generative grammar and psycholinguistic camps. C&C's experienced-based approach claims that the processing differences are simply due to "their relative distribution in the experience of individual language comprehenders" (p. 173). Supporting this conclusion, their analysis of a corpus of more than 11 million words in both spoken and written English found that examples of the first, subject-relative sentence occurred over twice as often as examples of the second, object-relative sentence. Thus, the argument is that simple experienced frequency of prior exposure is at least partially responsible for this difference. A nativist might simply reply that frequency of prior exposure is simply a performance influence and has little to do with understanding what is being said (which is the responsibility of generative grammar). But this reply seems much less important now because there is more agreement on what is responsible for actual performance.

More generally, language corpora are revolutionizing language inquiry, and their increasing ease of deployment (e.g., Language Goldmine, 2016) is winning many converts, me included. One persistent source of "evidence" that has long been central to the nativist claim that language is special is the poverty of the stimulus (Chomsky, 1980). The growing child simply does not hear enough language to account for her creative language use. Yet who would have thought that a typical child has probably heard millions of words during at least a thousand hours of speech before she reaches her terrible twos (Roy, Frank, DeCamp, Miller, & Roy, 2015). Strengthening the empiricist account is the recent finding that the vocabulary used during parental speech to children is highly correlated with the child's vocabulary (Massaro, 2016).

The authors also provide convincing evidence for general processes rather than language-unique processes in language understanding. This is also true, they claim, for people with speech language impairment (SLI). They propose that there is a plethora of perceptual/cognitive mechanisms at play in language processing, and some subset of these might account for SLI. For example, they show that poor sequence learning appears to account for some of the language processing deficits in SLI. More generally, it is possible that a deficit in general-purpose procedural learning might account for SLI and thus would also be apparent in nonlinguistic tasks.

A Personal Critique

Despite my admiration for the C&C book, the following critique is best understood by a caveat and full disclosure that I dedicated much of my research career to an outsider's study of language processing. I spent most of my career as one of the few voices against the dominant claim that speech is special. Early in my postdoc, I mentioned our speech research to David Green, the noted auditory psychophysicist, and encountered the reply, "Oh, that's very different." My dissident role had very little impact on the field, but this book and the plethora of research it reviews appear to substantiate the value of my early research trajectory. This value is only somewhat diminished by my neglect of evolutionary principles and neurological underpinnings of mind and behavior. With respect to the former, I argued that psychological explanation requires proximal influences on behavior regardless of the history of distal influences (Massaro, 1979). For the latter, I took a stance against the adequacy of a completely reductionist account of behavior and promoted the value of a more global functional account (Massaro, 1986).

Four decades ago, as a junior professor, I convinced a small cohort of graduate students mentored

_s
_n
_l

cs

by my new senior faculty colleagues to apply an information processing analysis to understanding language. Given the graduate students' areas of visual perception, verbal learning, and eyelid conditioning, and the fact that they resided in the Midwest, Chomsky's shadow did not exert much influence, and we could happily proceed with our quest. Our concern was to "view the understanding of language as a sequence of psychological (mental) processes that occur between the initial presentation of the language stimulus and the meaning in the mind of the language processor" (Massaro, 1975b, pp. 4–5). Central to this information processing framework, our primary concern was with real-time processing. The theoretical framework was grounded in structure and process. Memory structures constrain the processing that was possible. As an example, research indicated that the initial speech signal is stored in a preperceptual auditory store that lasts only about a quarter second. Some transformation is therefore necessary to create a more stable encoding (in this case, recognizing a so-called perceptual unit) that can be used by a succeeding stage of processing. This framework anticipated C&C's Now or Never Bottleneck (NNB) and Chunk-and-Pass perspectives, which assume that "the rich perceptual input is recoded as it arrives to capture the key elements of the sensory information as economically and distinctively as possible" (p. 97).

It is instructive to view our so-called information processing model juxtaposed with C&C's scheme (Figure 2). Constrained by the information processing framework, it seemed necessary to include both structural (memory) and functional (process) components to model how the language comprehender advances from the language input to understanding. C&C set themselves a somewhat more general charge of simply listing four increasingly abstract levels of processing.

Quantifying Language Processing

With their broad coverage, C&C tended to neglect important research findings on the speech side of language. This neglect is unwarranted: They do not have to fall into Hockett's trap of basically equating language with speech, but certainly speech is the primary materialization of the world's languages. They correctly claim that the sounds of speech are transient but give the duration of its initial sensory representation as less than 100 ms when it is best estimated at around a quarter of a second (Massaro, 1972). This larger estimate was based on auditory backward recognition masking (ABRM). Previously, research had focused on detection masking, in which a neighboring intense sound blocks hearing or detecting another soft sound. Whether the neighboring sound comes before, during, or after the target sound

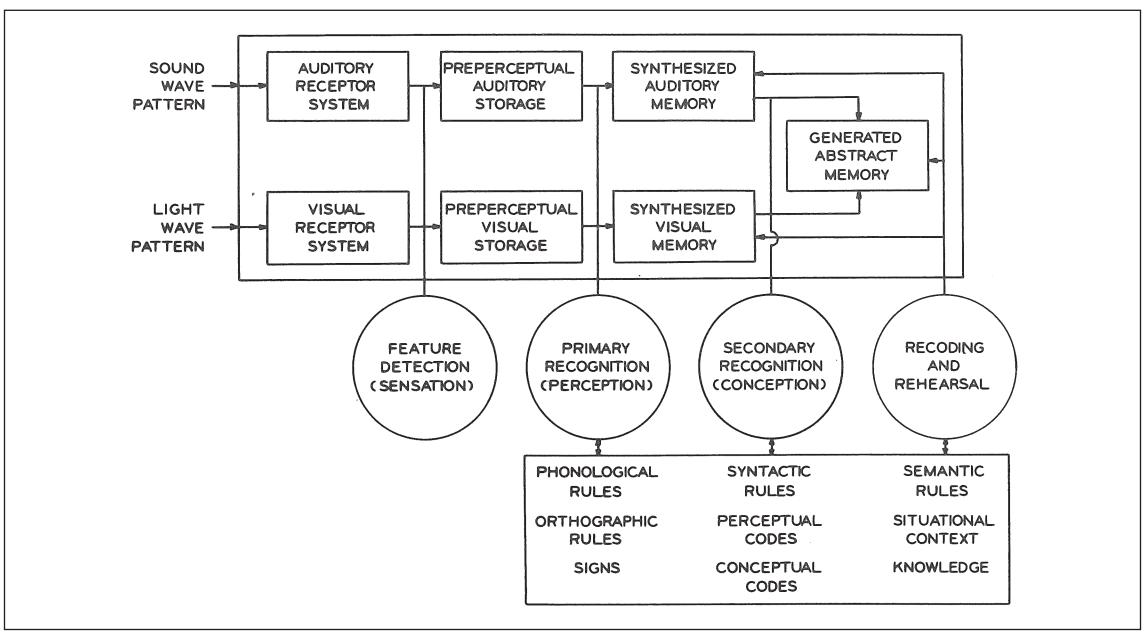

FIGURE 2A. Flow diagram of the temporal course of auditory and visual information processing as it appeared in Massaro, *Understanding Language: An Information Processing Analysis of Speech Perception, Reading and Psycholinguistics.* Academic Press, 1975. Used with permission. Elsevier, 2016

s_
n_
l_
cs

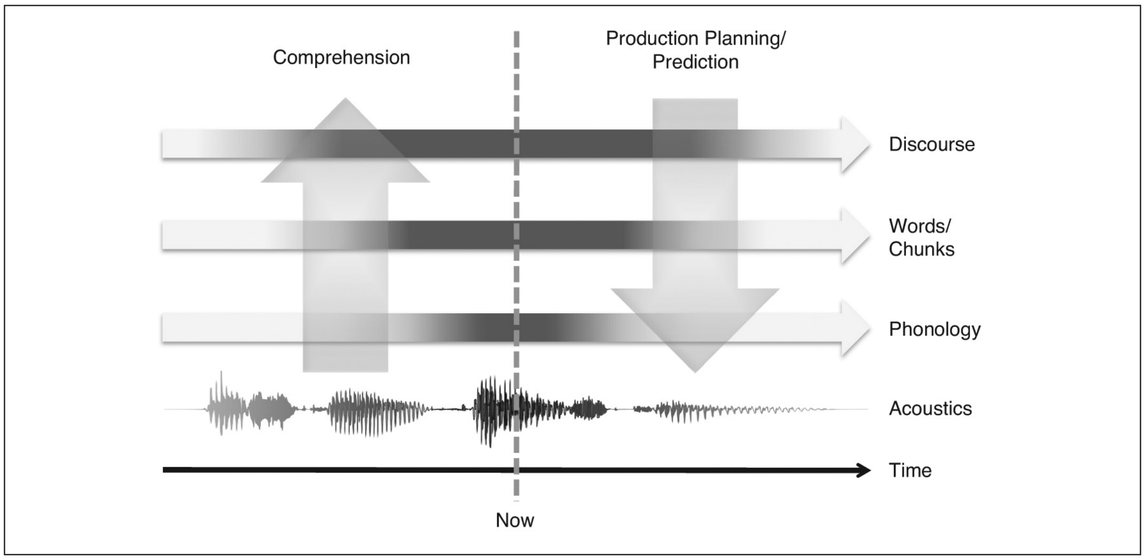

FIGURE 2B. The interrelations between the evolution, acquisition, and processing of language. From *Creating Language: Integrating Evolution, Acquisitions, and Processing*, by Morten H. Christiansen and Nick Chater, published by The MIT Press, ©MIT 2016 (used with permission)

is not critical. This type of detection experiment is the one that gives an answer of somewhere between 50 and 100 ms, which C&C use as their estimate. ABRM, on the other hand, does not interfere with detection because the neighboring sound is the same intensity as the target sound to be identified. Usually, the target sound varies on some auditory characteristic, such as the pitch or timbre of a sound, or is a speech category difference. Participants must identify the target, and the interference occurs only when the neighboring sound comes after, not before, the target sound. Interference can occur up to about 250 ms after the onset of the target, thus giving the larger estimate of the initial sensory representation in speech.

Why agonize about a difference of 150 ms or so? Given the shorter estimate, there would have to be as many as 10 transformations or recodings per second of speech. This number is excessively large and might easily overwhelm the processing system. Given the lack of invariance between the acoustic input and various phonemes, I used the larger duration to propose that a larger perceptual unit existed. This would not only bring the number of transformations per unit time into a reasonable range, but various research findings indicated the syllable unit could also restore a reasonable degree of constancy between it and auditory representation. I claimed V, CV, VC syllables, where V is a vowel and C is a consonant (or consonant cluster) as perceptual units for speech rather than the phoneme (Massaro, 1972, 2011). The phoneme has a clear lack of invariance between signal and percept, whereas these syllables restore most of the invariance needed for reliable pattern recognition (Massaro & Oden, 1980).

Meeting the Gavagai Challenge

The plethora of research literature the C&C review goes a long way to foster understanding how the child is easily capable of solving the Gavagai problem. There are many, many different sources of information that serve as potential constraints to allow the child to associate the appropriate meaning with the appropriate linguistic utterance. Given C&C's operating thesis of domain-general underpinnings of language use, I am troubled when they argue a case for distinguishing between the processes involved in understanding language and those used in pattern recognition in other domains. This distinction is between navigating the natural world (N-induction) and coordinating with our fellow interlocutors (C-induction). C&C state that "these two types of problems are very different" (p. 69). In N-induction we are measuring up to an immutable standard; in C-induction the standard is socially converted. If language has evolved to be learnable, the authors claim it may not present the same challenge of induction that typical pattern recognition involves. For objects, the perceiver has to induce what pattern is most likely (given the myriad of possible cues). For language, the perceiver simply has to determine the answer that the community of language learners has already agreed on. The authors make this argument because language is learned and learners have similar learning biases. These biases somehow give the learner an advantage because "the first wild guesses that the learner makes about how some linguistic structure works are likely to be the right guesses" (Christiansen & Chater, 2008, p. 507).

It is not apparent to me how the language task is any easier because in one situation the community of language users enforce the correct answer, and in the other the physical world does. Even if some first wild guesses were correct, the language learner would still have to learn and remember which were correct and which were not. In signal detection terminology, a bias does not improve sensitivity. The perceiver of both objects and language has to infer what has been encountered, based on the available information. I do not see how language evolved to facilitate pattern recognition more than the pattern recognizer evolving to make sense of fairly constant aspects of the physical world. Given that successive generations experience a fairly stable ecological niche but encounter a language that is continually changing, we might, in fact, actually expect more specialized innate constraints for N-induction than C-induction. Surely, the information will differ in the two domains, but I believe the information processing will follow the same format, inferring a meaningful event, whether it be some object input or some linguistic input. Our research, as well as that from many others, has demonstrated exactly analogous processes for understanding language and for recognizing objects and events (Massaro, 1998; Movellan & McClelland, 2001). Ironically, C&C's call for the apparent ease of language processing could almost be interpreted as a nativist argument, which the authors would abhor. Perhaps the authors might be convinced of this critique because they so strongly advocate fundamental perceptual, cognitive, and learning processes for language understanding as opposed to language-specific ones.

The Need for Formal Models of Language Understanding

If there is a downside to C&C's contribution, it is a plethora of discursive narrative without any formal models. Their chapter 4 on the NNB and Chunk-and-Pass processing might be read to allow many (sometimes) contradictory propositions. The main idea behind the NNB is that language input is highly transient, and it must be chunked and passed to a higher, more abstract level of representation. This is all very reasonable and reflects the progress speech and psycholinguistic science has contributed over the past decades. But the devil is in the details. Without a formal model as a guide, C&C might be interpreted to imply discrete categorization from level to level when they say, "The acoustic signal is first chunked into higher-level sound units at the phonological level. To avoid interference between local sound-based units, such as phonemes or syllables, these are recoded" (p 107). Just a few pages later, however, they say their proposal "fits with proposals . . . where local ambiguity resolution is temporally delayed until later disambiguating information arrives" (p. 113). Given these statements, the authors do not take a strong stance on what type of processing and recoding occur at each increasing abstract level. It is well known, and recognized by the authors, that later information can resolve ambiguity in earlier-arriving information. But for one to take advantage of the later information, the previous recoded information cannot be discrete or categorical. If language perceivers are to benefit from two sources of information, the sources must be graded and not discrete (Massaro, 1987).

Like the Gavagai challenge, language understanding might appear to be an insurmountable problem, which is perhaps why Chomsky was so successful in convincing the field of a nativist solution. According to C&C, the high quality of language processing follows from the use of multiple constraints. They propose that the perceiver performs parallel integration of multiple cues at multiple levels of language processing. Now that Bayesian reasoning has been featured in cognitive science, their solution is palatable to most of the current players. There is now much convincing evidence of a Bayesian type integration in speech perception and reading (Massaro, 1998; Massaro & Jesse, 2005). What is ironic, however, is that the authors do not review a single study illustrating Bayesian-like integration in language processing. Their elegant corpus analyses undoubtedly reveal the ecological validity of multiple constraints or cues in typical spoken and written language. However, they do not succeed in describing experiments demonstrating that multiple cues are actually used together to facilitate language processing.

C&C repeatedly postulate that multiple cues are integrated; chapter 5 actually includes "Multiple-Cue Integration" in its title. I have defined several possibilities of how multiple cues could be used and

s_
n_
l_

cs

formally operationalized integration as the simultaneous use of two or more cues to categorize a single presentation of a language event (Massaro, 1987). One of the most popular illustrations of this type of integration is when the sound of speech and the facial movements of the talker are used together to identify a speech syllable (the so-called McGurk effect). Our research has taken great pains to distinguish between the many ways multiple sources of information can be used. So, for example, using the most informative cue on each trial would not be equivalent to an optimal Bayesian-type integration. Within our fuzzy logical model of perception (FLMP), each cue is assigned a truth value indicating the degree to which it supports each potential categorization. These truth values supporting a given categorization are then multiplied and evaluated against all other possible categorizations. Truth values are a good metric because they more naturally represent graded information compared with probabilities that can easily be interpreted as discrete (Massaro, 1998). As formulated, the FLMP is mathematically equivalent to Bayes's theorem, which can be interpreted as an optimal use of multiple cues.

When we began our language studies, the predominant experimental strategy (except for a few notable exceptions) was to manipulate only a single source. I advocated the approach of manipulating several sources of information independently of one another in language processing tasks (Massaro, 1975a). This paradigm was necessary because manipulating just a single source would not illuminate how that source interacted with other sources. In addition, by neutralizing other sources in a single-source experiment, participants might easily zero in on that source even though they do not normally use it in a productive manner. Manipulating multiple sources, for example, we evaluated how both the letter quality and the frequency of orthographic patterns influenced letter and word recognition in both speech and reading domains. Based on our research and that of others, we have also claimed that the robustness of language processing results from the efficient use of multiple top-down and bottom-up sources of information (Massaro, 1979).

C&C discuss top-down and bottom-up sources of information when they describe how both contextual and acoustic information are used to identify word recognition. To rationalize the evolution and development of an arbitrary relationship between the form of a word and its meaning, they cite mathematical evidence that two sources of information provide maximal constraint when they are independent of one another. Thus, they propose that multiple cues do not help if they are redundant. In language processing, as in other forms of pattern recognition, however, their redundancy is a necessary condition. In auditory/visual perception, for example, the two sources of information are necessarily redundant because they come from the same speech utterance. Similarly, constraining context will facilitate word processing even when the form of the word also predicts its meaning. The reason is that perceiver treats these two sources as mostly independent of one another (Massaro & Stork, 1998) and benefits from having two evaluations relative to just one. The perceiver gets partial information from each cue, and a Bayesian type integration provides more information given both cues rather than just one. Using this framework, we have provided a good quantitative description of how these two sources of information are integrated in both speech perception and reading (Massaro, 2012b).

Our research in auditory–visual bimodal speech perception not only demonstrates the value of two sources of information relative to just one, it also adds convincing support for experience in language processing. Adults and children of various ages identified speech syllables with either consistent or inconsistent auditory and visual properties. All participants used both modalities, but the younger the children, the less they benefited from the visible speech. Speech perception and language more generally are slowly acquired skills and consistent with C&C's claim that language acquisition can be interpreted as learning to process language. We make a further distinction between information and information processing, operationalized as the informativeness of a source of information and how the sources are used together. Clearly, informativeness increases with experience, but there is evidence for an optimal integration of sources in auditory/visual speech perception across ages 3 to 83 (Massaro, 1998).

C&C proselytize the benefit of having multiple sources of information. However, having two discrete sources of information would not benefit a correct resolution of the linguistic input. If the two sources agreed, there could not be any advantage to having only one source. If the two sources disagree, they would be no information to guide which source should be followed. Thus continuous or graded encodings are necessary to assuming that graded information from each source was one of the central assumptions of the FLMP to predict the integration of top-down and bottom-up sources of information in both speech perception and reading (Massaro, 1979; Massaro & Oden, 1980). The framework of the FLMP provides not only a coherent description but a testable quantitative one.

The FLMP is grounded in fuzzy logic in which a proposition has a degree of truth rather than just true or false. The NNB and the Chunk-and-Pass operation could therefore be formalized within the FLMP framework to pass continuous rather than categorical information from one level to the next more abstract level. To illustrate the FLMP and how it is tested in experiments, consider how Massaro and Oden (1995) analyzed Pitt's (1995) study of the joint influence of phonological information and lexical context in an experimental paradigm developed by Ganong (1980). A speech continuum was made between two alternative CVC syllables *gift* and *kiss,* and the contextual information was varied to support one alternative or the other. The initial consonant of the CVC syllable was varied in six steps between /g/ and /k/. The following context was either /ift/ or /is/. The context /ift/ favors or supports initial /g/ because *gift* is a word, whereas *kift* is not. Similarly, the context /is/ favors or supports initial /k/ because *kiss* is a word, whereas *giss* is not. Pitt improved on earlier studies by collecting enough observations to allow a subject-by-subject evaluation of the ability of specific models of language processing to account for the results. Previous tests of models using this task have depended primarily on group averages, which may not be representative of the individuals that make the averages up. The results showed a strong effect of both the initial consonant and the following context and a significant interaction between these two variables.

According to the perceptual unit assumption, the CVC is composed of two successive units. Applying the FLMP, it is assumed that the initial CV is not categorically perceived but rather is represented as the degree to which it corresponds to a /gi/ or /ki/. This information is passed on to the next level of word identification, given the additional source of information /ift/ or /is/. These two sources of information are combined to give an overall degree of support for the words *gift* or *kiss*. The FLMP provided a good quantitative fit for each of the 12 subjects in the task. Assuming independence of bottom-up and top-down sources of information is critical for a coherent description of their joint influence (Massaro & Cohen, 1991). Later-arriving context does not change the representation of the earlier-arriving bottom-up information. The updated information is simply represented at a more abstract level. Thus, in the *gift–kiss* example, the two sources are combined at the word level, not at the initial CV syllable level. Evidence for this independence was found in a signal detection analysis revealing that the information about the bottom-up source remained independent of the top-down source (Massaro & Oden, 1995).

In summary, several assumptions are necessary to account for research findings within the NNB and Chunk-and-Pass framework. First, categorizations at each level must be graded rather than discrete to take advantage of multiple sources of information at multiple levels. Second, new information does not feed down and modify the representation or categorization of an earlier level. The new information as in the word ending in the *gift–kiss* experiment simply functions as an additional independent source of information. Of course, the word recognition can also be graded as substantiated by continuous rating judgments. This form of processing also describes how the perceiver traverses through more cognitively complex language understanding.

Form Meaning Correspondences

C&C provide a thorough treatment of the issue of whether the correspondence between the form of a spoken word and its meaning is arbitrary or systematic in some way. They define absolute and relative iconicity. Absolute iconicity occurs when a linguistic feature directly imitates some semantic characteristic. Thus, *choochoo* represents the sound that a train makes (or at least used to make). These words are usually described as onomatopoeic words. However, Perry, Perlman, and Lupyan (2015) have established a role for absolute iconicity beyond obvious imitation, which they define as a cross-modal correspondence in the analog properties of a word's meaning and its spoken symbol. Measures of iconicity can be derived from adults' judgments of how much English words sound like what they mean. Onomatopoeic words such as *slurp* would be rated as highly iconic. Words such as *teeny* and *huge* are iconic because they sound small and big, respectively. Words such as *cat* and *dog* do not appear to have any cross-modal correspondence between their sound and meaning.

C&C define relative iconicity as a statistical regularity between sounds and meanings in the absence of imitation (p. 139; see also Monaghan, Shillcock, Christiansen, & Kirby, 2014). They carried out a series of corpus analyses to explore the role of relative iconicity in vocabulary and syntax. Although statistically significant relative to chance, the amount of variance accounted for by sound–meaning mappings in this case was very small, with less than 0.2% of the variance accounted for. To find evidence for a role of iconicity, Monaghan et al. (2014) continued to find significant correlations even when all monosyllabic words that shared morphophonic and etymological roots were eliminated from the analysis. Moreover, even if these morphophonic and etymological roots shared between words are completely accounted for, there may be other constraints besides iconicity that are contributing to this correlation. Although the authors appear to use the terms *systematicity* and *relative iconicity* interchangeably, I suggest that the term *iconicity* be reserved for a cross-modal correspondence between form and meaning that can be rationalized as having some imitative property. As observed by Winter et al. (in press), the English cluster *gl* bears no obvious resemblance to the meaning of shiny visual phenomena, but it repeatedly occurs in words such as *glitter, glimmer,* and *glisten* (Winter et al., in press).

It is reasonable to expect that iconicity would play a larger role early in language acquisition than later, when the vocabulary would necessarily become more arbitrary as it increases in size. Like other investigators, C&C use adults' estimated age of acquisition (AoA) as a measure of the age at which words are acquired. Although one might think that judging the age when specific words were learned would be unreliable, research has shown that AoA ratings and overall frequency occurrence have very large and roughly equivalent correlations with reaction time and accuracy in a lexical decision task (Kuperman, Stadthagen-Gonzalez, & Brysbaert, 2012; Kuperman & Van Dyke, 2013). However, C&C attempt to make the case for a larger influence of their measure of relative iconicity on AoA estimates at ages 2 and 3 than at later ages, but the results shown in their Figure 5.2 are not very convincing. Other results, on the other hand, have found evidence for absolute iconicity influencing early vocabulary learning (Perry et al., 2015).

Continuing to explore systematicity in English, C&C's corpus analyses also evaluated the phonological properties of noun and verb categories, and they found that these properties could predict the two categories above chance. They then determined distributional properties of the words by quantifying the likelihood of a word following one of 20 most frequent words in the corpus. This cue also predicted noun–verb category membership well above chance. They then claim to show that with the simple combination of the phonological and distribution cues "good classification can be found" (p. 153). However, the claimed synergy that the authors show in their Figure 5.3 is not really there. Performance given both sources is never better than performance given just the most informative source. Their method of simply representing the sources in a multidimensional space might be responsible if they do use an appropriate integration algorithm (Massaro & Friedman, 1990).

This is a productive research program documenting the possible multiple cues supporting language acquisition and processing. As recognized by the authors, uncovering predictive properties in the language does not address how these properties are actually processed by the language perceiver. It is important for researchers to keep in mind a distinction between ecological and functional validity in terms of whether an ecological cue is actually functional in language processing (Massaro, 1987).

Retrospective

C&C's last chapter builds a strong foundation for integration rather than fragmentation in language studies. Integration corresponds to language disciplines as well as behavioral principles that are not unique to language. The present volume certainly reinforces Skinner's original faith in the environment's influence on language processing. It remains to be seen whether their general framework will supplant an alternative nativist account. Like the last decade of the question of climate change, we might have to accept a prolonged process mimicking evolution itself.

Note

I thank Marcus Perlman, Robert Proctor, and Bill Rowe for reading and commenting on the review.

Dom Massaro
Department of Psychology
Social Sciences II
1156 High Street
University of California–Santa Cruz
Santa Cruz, CA 95064
E-mail: massaro@ucsc.edu

REFERENCES

Anderson, M. L. (2010). Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences, 34,* 245–266.

Chomsky, N. (1959). A review of B. F. Skinner's *Verbal Behavior*. *Language, 35*(1), 26–58.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1980). *Rules and representations*. Oxford, England: Basil Blackwell.

_s
_n
_l
cs

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral & Brain Sciences, 31,* 489–558.
Everett, D. L. (2005). Cultural constraints on grammar and cognition in Pirahã. *Current Anthropology, 46*(4), 621–646.
Freund, A. (1975). Word and phrase recognition in speech processing. In D. W. Massaro (Ed.), *Understanding language: An information processing analysis of speech perception, reading and psycholinguistics* (pp. 357–389). New York, NY: Academic Press.
Ganong, W. F. III. (1980). Phonetic categorization in auditory word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 6,* 110–125.
Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press.
Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods, 44*(4), 978–990. Retrieved from http://doi.org/10.3758/s13428-012-0210-4
Kuperman, V., & Van Dyke, J. A. (2013). Reassessing word frequency as a determinant of word recognition for skilled and unskilled readers. *Journal of Experimental Psychology: Human Perception and Performance, 39,* 802–823.
Language Goldmine. (2016). Retrieved from http://language goldmine.com/
Massaro, D. W. (1972). Preperceptual images, processing time, and perceptual units in auditory perception. *Psychological Review, 79,* 124–145.
Massaro, D. W. (1975a). *Experimental psychology and information processing*. Chicago, IL: Rand McNally.
Massaro, D. W. (1975b). *Understanding language: An information processing analysis of speech perception, reading and psycholinguistics*. New York, NY: Academic Press.
Massaro, D. W. (1979). Reading and listening (tutorial paper). In P. A. Kolers, M. E. Wrolstad, & H. Bouma (Eds.), *Processing of visible language* (Vol. 1, pp. 330–354). New York, NY: Plenum.
Massaro, D. W. (1986). The computer as a metaphor for psychological inquiry: Considerations and recommendations. *Behavior Research Methods, Instruments, & Computers, 18,* 73–92.
Massaro, D. W. (1987). *Speech perception by ear and eye: A paradigm for psychological inquiry*. Hillsdale, NJ: Erlbaum.
Massaro, D. W. (1998). Perceiving talking faces: From speech perception to a behavioral principle. Cambridge, MA: MIT Press.
Massaro, D. W. (2011). *Reflections on the syllable as the perceptual unit in speech perception*. Retrieved from http://www.talkingbrains.org/2011/02/reflections-on-syllable-as-perceptual.html
Massaro, D. W. (2012a). Acquiring literacy naturally: Behavioral science and technology could empower preschool children to learn to read naturally without instruction. *American Scientist, 100,* 324–333.
Massaro, D. W. (2012b). Speech perception and reading: Two parallel modes of understanding language and implications for acquiring literacy naturally. *American Journal of Psychology, 125,* 307–320.
Massaro, D. W. (2015, May 12). Method and system for acquisition of literacy. United States patent number U.S. 9,028,255 B2.
Massaro, D. W. (2016). Multiple influences in vocabulary acquisition: Parental input dominates. In *Proceedings of the 17th Annual Conference of the International Speech Communication Association (Interspeech 2016),* 878–882. Available at www.isca-speech/archive/interspeech_2016/pdfs/0037
Massaro, D. W., & Cohen, M. M. (1991). Integration versus interactive activation: The joint influence of stimulus and context in perception. *Cognitive Psychology, 23,* 558–614.
Massaro, D. W., & Friedman, D. (1990). Models of integration given multiple sources of information. *Psychological Review, 97,* 225–252.
Massaro, D. W., & Jesse, A. (2005). The magic of reading: Too many influences for quick and easy explanations. In T. Trabasso, J. Sabatini, D. W. Massaro, & R. C. Calfee (Eds.), *From orthography to pedagogy: Essays in honor of Richard L. Venezky* (pp. 37–61). Mahwah, NJ: Erlbaum.
Massaro, D. W., & Oden, G. C. (1980). Speech perception: A framework for research and theory. In N. J. Lass (Ed.), *Speech and language: Advances in basic research and practice* (Vol. 3, pp. 129–165). New York, NY: Academic Press.
Massaro, D. W., & Oden, G. C. (1995). Independence of lexical context and phonological information in speech perception. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1053–1064.
Massaro, D. W., & Stork, D. G. (1998). Sensory integration and speechreading by humans and machines. *American Scientist, 86,* 236–244.
Monaghan, P., Shillcock, R. C., Christiansen, M. H., & Kirby, S. (2014). How arbitrary is language? *Philosophical Transactions of the Royal Society B: Biological Sciences, 369,* 20130299.
Movellan, J., & McClelland, J. L. (2001). The Morton–Massaro law of information integration: Implications for models of perception. *Psychological Review, 108,* 113–148.
Perry, L. K., Perlman, M., & Lupyan, G. (2015). Iconicity in English and Spanish and its relation to lexical category and age of acquisition. *PLOS ONE, 10*(9), e0137147.
Pitt, M. A. (1995). The locus of the lexical shift in phoneme identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1037–1052.
Quine, W. V. O. (1960). *Word and object: An inquiry into the linguistic mechanisms of objective reference*. Cambridge, MA: MIT Press.
Quine, W. V. O. (1990/1992). *Pursuit of truth.* Cambridge, MA: Harvard University Press.
Rizley, R. C., & Rescorla, R. A. (1972). Associations in second-order conditioning and sensory preconditioning.

s_
n_
l_
cs

Journal of Comparative and Physiological Psychology, 81, 1–11.
Roy, B. C., Frank, M. C., DeCamp, P., Miller, M., & Roy, D. (2015). Predicting the birth of a spoken word. *Proceedings of the National Academy of Sciences of the United States of America* (PNAS). Retrieved from http://www.pnas.org/content/early/2015/09/15/1419773112
Skinner, B. F. (1957). *Verbal behavior.* Acton, MA: Copley Publishing Group.
Wang, T., & Saffran, J. R. (2014). Statistical learning of a tonal language: The influence of bilingualism and previous linguistic experience. *Frontiers in Psychology, 5,* 953. doi:10.3389/fpsyg.2014.00953
Winter, B., Perlman, M., Perry, L. K., & Lupyan, G. (in press). Which words are most iconic? Iconicity in English sensory words. *Interaction Studies*.

RETHINKING AUTISM'S PAST, PRESENT, AND FUTURE

Neurotribes: The Legacy of Autism and the Future of Neurodiversity

By Steve Silberman. New York, NY: Avery, 2015. 534 pp. Hardcover, $29.95

> ***Neurotypical syndrome is a neurobiological disorder characterized by preoccupation with social concerns, delusions of superiority, and obsession with conformity. There is no known cure.***
> ***—Institute for the Study of the Neurologically Typical (Silberman, 2015, p. 441)***

The legacy of autism is not a happy one. Throughout history, autistic people have been misunderstood and marginalized, isolated and ignored. They have been bullied by peers, abandoned by parents, and murdered by the Nazis. They have been the subject of and subjected to some of the worst treatments (and science) imaginable, involving everything from chelation to electric shocks. And that is just in the last century.

For readers who have only a passing acquaintance with autism, a brief primer may be helpful. Autism is a lifelong neurological condition whose causes are unknown. There is no biological marker for autism, no blood test, no brain scan; it is diagnosed on the basis of clinical observation. The criteria in the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders* involve impairments in communication and social interaction and the presence of restrictive and repetitive patterns of behaviors, interests, or activities (American Psychiatric Association, 2013). These symptoms must be present from early childhood, must cause clinically significant impairment, and cannot be better explained by intellectual disability. Recent estimates in the United States suggest autism affects about 1 in 68 people (Centers for Disease Control, 2014), although there are tremendous differences in prevalence rates across countries (Norbury & Sparks, 2013).

One of the things that makes autism so fascinating and frustrating is its heterogeneity. Some autistics talk too much, some do not talk at all. Some score off the charts on IQ tests, some are unable to take those tests. Some cannot stand bright lights or loud sounds, some thrive on them. There is a saying in the autism community that if you know one person with autism, you know one person with autism. There is no cure for autism. The data on outcomes for adults are paltry, but what few data exist suggest that outcomes are terrible: A study on autistic adults in their 40s (who were diagnosed as children and had average nonverbal IQ scores at that time) found that most did not live independently, were unemployed, and had never had a meaningful relationship (Howlin, Moss, Savage, & Rutter, 2013).

The good news, according to science journalist Steve Silberman's engrossing new book *Neurotribes: The Legacy of Autism and the Future of Neurodiversity,* is that a movement is afoot, one that may fundamentally change for the better how autistic people are treated, studied, and educated. Neurodiversity advocates hold the radical view that autism and other neurological differences are not devastating disorders in need of curing or eliminating. They are instead part of the natural variation of the human condition—variation that can result in unique challenges, to be sure, but also in unique strengths. The reason autism can be so debilitating, the argument goes, has more to do with society's lack of support, accommodation, and understanding than with autistics' atypical neurology. We should be working to create inclusive communities where autistics can flourish with (and perhaps because of) their autism, not trying to turn autistics into nonautistics.

The neurodiversity movement is young (the term was coined in the 1990s), but the ironic thesis of Silberman's book is that "viewing [autism] as a lifelong disability that deserves support, rather than as a disease of children that can be cured . . . [is] the oldest idea in autism research" (p. 81). According to Silberman, the person responsible for that idea and the hero of his story is Hans Asperger (1906–1980), a pediatrician who directed the special education unit at the University of Vienna's Children's Clinic. In