The Relevance of Chomsky in 21st Century Second Language Acquisition

James Heather

Abstract

As the world continues to evolve so do theories of how humans learn language. The ideas proposed by linguist Noam Chomsky in the second half of the 20th century continue to linger today. The post-Chomskian thrust views language acquisition in terms of constructivism. Constructivism implies that learning language is less about Chomsky’s ideas of innate grammar, studying the final state or language acquisition devices and more about analyzing acquisition processes. This paper summarizes some of Chomsky’s ideas, subjects them to critical analysis, before looking at the current movement of alternative theories and considering their contributions to the current and future direction of Teaching English as a Second/Foreign Language (TESL/TEFL). The paper concludes the current constructionist thrust represents a move away from Chomsky’s theoretical dominance and suggests the future of applied linguistics is on the verge of another new frontier.

Keywords: applied linguistics, constructionism, post-Chomsky, TESL, TEFL

“For all its empirical weaknesses Chomsky and generative grammar has ruled the linguistic world for 40 years” (Tomasello, 2003)

The world has changed greatly since the end of the Second World War. Thirteen United States presidents have served in office, a man has been put on the moon, medical, biological, and psychological technologies have advanced in leaps and bounds, and computers, the Internet, and social media have become essential everyday items. The list of modern-day changes in human existence
goes on. When considering the grand scale of these societal changes, having the same linguistic theory over fifty years roughly indicates one of two things. Either the theory is absolutely flawless and correct, or the academic world of linguistics has reached an impasse. Considering the time span, it would perhaps be more accurate to suggest the impasse occurred a while back, and that only a minority of academics are posturing and toying around the frontiers of new, revamped, or extended second language acquisition (SLA) theories. It is beyond the scope of this paper to speculate as to why it is taking so long for the linguistic world to move towards new theories. However, with the writings of the likes of Pinker (1994, 1995), Searle (2002), Tomasello (2003), Ellis (2003), Deacon (1997) and others, it is worth questioning the relevance of Chomsky’s theories in the 21st-century world of English as a Second Language (ESL). In this paper I take on the challenge of summarizing some of Chomsky’s ideas and how they have influenced Teaching English as a Second Language (TESL). After that I subject these ideas to critical analysis, before finally looking at the current thrust of alternative theories and considering their contributions to the current and future direction of TESL.

**Chomsky Influence**

It is not easy to summarize the full range of Chomsky’s ideas. Some have changed over the years, and his theory of linguistics has evolved over time from syntactic structures to Universal Grammar (UG) to the Minimalist Program. Chomsky himself has fine-tuned his theory over the years. While the characterization of UG as genetics was the stated goal of contemporary linguistics almost from its creation (Chomsky, 1965), empirical research had very little to contribute until the early 1980s when Chomsky notes: “the first major breakthrough took place with the principles and parameters framework ... when the concepts begin to be realized in terms of concrete empirical research” (2000, p. 8). It has been argued that the earlier lack of empirical research is attributable to the fact that the design principles of UG were largely unknown in detail. Seemingly, there were some aspects of the model that only Chomsky himself could explain. His ideas seem to provide answers to some of our deepest puzzles about language and the mind. Some of Chomsky’s ideas are deep, complex, and arguably difficult, and cannot be summarized briefly. The focus in this paper is on
the relatively limited field of Chomsky’s ideas that have influenced TESL practice, from the various periods of his career.

The theory originally proposed by Chomsky argued that human language had to be innate and intuitive. Chomsky proposed that a child is programmed to learn language, and that humans are born with a Universal Grammar, a set of rules or a basis for human language. Of UG, Chomsky said it is “the system of principles, conditions and rules that are elements or properties of all human language ... the essence of human language” and “UG is a theory of knowledge; its concern is with the internal structure of the human mind” (1976, p. 29). From this Chomsky implied that UG is wired in our brains, specifically through what was described as a Language Acquisition Device (LAD). The LAD set the parameters for grammar in which a child first recognizes what kind of language the parents are dealing with, and then sets his grammar to the correct one. Children, he says, master grammar unconsciously. The innate form of language is Universal Grammar, which, according to this theory, is common to all languages. It consists of deep structures and surface structures as well as a natural order of acquisition in which we acquire these structures in a particular order, much like computer software wired to the brain.

Chomsky never proposed any approach to language teaching, let alone second language teaching. The methods that have evolved based on his theories have come from within the applied linguistics teaching community of which Chomsky (who would no doubt agree) is no longer a subscribing member. Worthy of inclusion here is the fact that Chomsky’s idea of innatism is one of many of the origins of language that make up the array of today’s teaching methodology. Included on this list are creationist theories, evolutionary theories, Chomsky’s nativist theory, the evolutionary nativist theory which include the likes of Pinker (1994) and the language instinct and the opposing theory of behaviorism, which is enjoying somewhat of a neo-revival in the form of connectionism. The last point of note concerns the debate over the influence of age when learning. Krashen et al. (1975) believe there are no great differences in how a child or an adult learns the language: that is, they learn through unconscious acquisition. The opposing view suggests that children and adults learn in different ways. These differing perspectives, along with the many theories of the origins of language, represent the dynamic pool of information from which TESL methodology evolves.
Chomsky’s ideas of innatism consider language acquisition to be a subconscious process. Language teaching methodology in the mid-20th-century adapted to accommodate for this point of view by gradually moving away from behaviorist-style learning and grammar analysis and emphasizing acquisition activities and communicative language teaching. Arguably, a connection may be made that Chomskian ideas have played a part in the development of CLT. This is a good thing. Chomsky’s ideas also had a great influence in SLA with Krashen and his Monitor Model and Natural Approach (Krashen & Terrell, 1983). While this approach may not have taken root in language teaching methodologies as profoundly as CLT it still nonetheless lives on in the minds of curriculum makers who support more humanistic methods.

**Chomsky Critics**

Tomasello (2003) recalls how Pinker, in 1984, expressed disappointment in the field of language acquisition in children. The hope in the 1960s that Chomsky’s UG would lead the way in describing the origins of language didn’t gain the steam of progression as anticipated. In fact, the pace of progression arguably lagged. Pinker attributed the lack of progress to “a lack of linguistic sophistication in the field” (in Tomasello, 2003, p. 323). Pinker’s analysis concurred with Chomsky in that his work accepts that human language is innate and programmed in the genes. However, Pinker and his supporters suggested that Chomsky was wrong about evolution. Pinker later described his evolutionary model in *The Language Instinct* (1994). In it he argued that the language instinct evolved in all its complexity step by step over hundreds of generations just as the human eye or an elephant’s trunk has evolved. Chomsky too, no longer in denial, described an evolutionary model for the language instinct (Chomsky, Hauser, & Fitch, 2002). In addition to this, geneticists have contributed to this debate in ways unthinkable not so long ago. Cavalli-Sforza (2000), for example, described how DNA could establish the genetic connections among all the populations of the world. All of this has led to a line of theory that supports an evolutionary-nativist approach, an alliance of innatism and Pinker’s theory of biogenetics.

Deacon (1997) believes that artificial intelligence (AI) is a very real concept. In his book *The Symbolic Species*, he rejects the idea of UG and innate linguistic
knowledge and instead believes that language and the brain co-evolved: that is, 
the evolution of the human brain was driven by language mainly through the 
development of symbolic references (1997, p. 87). To Deacon, language did 
not require the emergence of a language organ. Rather, he paints an interesting 
portrait of the human race when he suggests that language originated from 
symbolic thinking and attributes this innovation to the time when humans became 
hunters because of the need to overcome sexual bonding in favor of group 
cooperation (see Scaruffi, 1999). He claims “Chomsky’s universal grammar is an 
outcome of the evolution of language in our mind during our childhood. There is 
no universal grammar in our genes ... there are no language genes in our genome” 
(Deacon in Scaruffi, 1999, p. 1). This point of view also counters Pinker’s 
assertion of a relationship between genetics and language.

In Searle’s harsh review of Chomsky’s New Horizons in the Study of Language and 
Mind (2000), he concurs with Pinker when he says that after years of the “Chomsky 
Revolution” the results are inconclusive (Searle, 2002). He claims Chomsky 
has altered or abandoned his original theory to a point where it contradicts 
itself. Searle comments: “Chomsky insists that the study of language is a branch 
of natural science” but that Chomsky’s newly revived theory of a language 
includes a lexicon plus computations. Searle’s objection arises from a belief that 
“computation is not a natural science ... but an abstract mathematical notion” 
(2002).

**Post-Chomsky Thrust**

The post-Chomskian thrust really began to take momentum in the 1990s. Ellis (2003) wrote a seminal article that summed up the progress of the previous 
15 years. Ellis views the post-Chomskian thrust in terms of constructivism. 
Within this “tribe” of theories he includes connectionists, functional linguistics, 
emergentists, cognitive linguistics, constructionist child language researchers, applied 
linguists influenced by chaos theory, and computational linguists. Ellis believes 
that all of these tribes “share a functional-developmental, usage-based perspective 
on language” (2003, p. 63). For Ellis, constructionist views of language acquisition 
hold that “simple learning mechanisms operating in and across human systems 
for perception, motor action and cognition while exposed to language data in a
communicatively rich human social environment navigated by an organism eager to exploit the functionality of language are sufficient to drive the emergence of complex language representations” (2003, p. 63). The most compelling aspect of Ellis’s writing is his belief that the structural regularities of a language “emerge from a learner’s lifetime analysis of the distributional characteristics of the language” (2003, p. 64). By this he implies that it is not a matter innate grammar as Chomsky believes, but rather a “statistical ensemble of language experiences that changes slightly every time a new utterance is processed” (2003, p. 64). Ellis suggests that the acquisition processes need to be analyzed rather than the final state or the language acquisition device. Constructivists, he says, “work within the broad remit of cognitive science, seeking functional and neurobiological descriptions of the learning processes which, through exposure to representative experience, result in change, development and the emergence of linguistic representations” (2003, p. 64). This counters Chomsky’s belief that “no discipline can concern itself in a productive way with the acquisition and utilization of a form of knowledge without being concerned with the nature of that system of knowledge” (1977, p. 43).

**Cognitive Linguistics**

The Nature-Nurture debate in science has been argued back and forth for many decades. Is it our genes or our environment that determines our personality and behavior? The debate stretches, overlaps, and is intertwined into other disciplines including psychology and applied linguistics. Is learning a first language related to evolutionary-nativist theory or some sort of neo-behaviorist (non-nativist) theory? To what extent can (or should) TESL instructors use this information regarding first language acquisition and apply it to SLA? There is an abundance of theories out there regarding how humans acquire a first language. This paper has focused on Chomsky’s ideas of innatism in acquiring language and has discussed the contributions/implications to TESL methods. In addition it introduced other positions and compared them in relation to Chomsky. For all of the “newness” these theories claim, the reality is they may be seen as mere extensions of the original nativist/non-nativist (nature/nurture) perspectives. The addition of “cognitive” theories has confused things unnecessarily. What is cognitive? If we view cognitive as being related to the conscious and/or unconscious brain/
mind then one could easily argue that Chomsky’s theory is cognitive. Until the late 1980s it could be argued that cognitive theories in language acquisition have contributed to the confusion and lack of progress in our field (i.e., the symbolic models as described in Gasser, 1990). Without clear universal understanding of what is meant by cognitive, it is easy to suggest that both nativist/non-nativist (nature/nurture) theories contain elements of cognitivism. More recently, cognitive linguistics has been said to “provide detailed qualitative analyses of the ways in which language is grounded in human experience and in human embodiment, which represents the world in a very particular way. The meaning of the words of a given language, and how they can be used in combination depends on the perception and categorization of the real world around us” (Ellis 2003, p. 65). Cognitive linguistics then, is said to reflect and embody the experience of everything we know, and everything we perceive.

**Constructivism**

Tomasello has been labeled a constructivist of the “child language research” variety by Ellis. Tomasello critiques UG by viewing it as “acquisition processes [that] connect somehow with an innate universal grammar.” He goes on to point out: “but these extra processes are completely unnecessary – important to save a theory but not to explain the phenomenon” (2003, p. i). The resonance of this statement could be fatal to some teaching methodologies. Chomsky’s ideas triggered such humanistic approaches to language teaching such as Suggestopedia (see Lozanov, 1978) or Counseling Learning (Curran, 1976). Without Chomsky’s theories these methods would scantily be justified, causing potential havoc for the teachers who adopt these humanistic approaches. Tomasello takes aim at Chomsky in much the same way as Pinker did. However, Tomasello goes one step further and also takes aim at Pinker and claims he used adult analyzing tools (lexical functional grammar) to discover how children become skilled users of a language. Tomasello suggests Pinker had to “make a continuity assumption: underneath, the language of children is structured by the same abstract categories and rules of adults” (2003, p. 323). This, Tomasello declares, “was a mistake” (2003, p. 324). For Tomasello, nativist theories miss the mark. He considers the recent advancements of developmental psychology, linguistics, and cognitive science and concludes: “children can get from here to there ... without the aid
of any hypothesized universal grammar” (2003, p. 326). Tomasello outlines two skills as being particularly important for language acquisition: skills of intention-reading and the skills involved in various kinds of pattern-finding-categorization. He argues that the essence of language is its symbolic dimension, which rests on the uniquely human ability to comprehend intention. For Tomasello, Grammar emerges as the speakers create language out of recurring sequences of symbols. Children pick up these patterns in the conversations they hear around them. Tomasello also critiques connectionism as a theory that has much to offer, however, at the moment he claims the models are “... psychologically unrealistic in two basic ways. First they do not deal with communicative intentions or function [and] second, connectionist models [currently] work only with very small units such as words and grammatical morphemes” (2003, p. 24). Tomasello offers a usage-based approach as an alternative to UG and connectionist models. First, his model is functionalist: “based explicitly in the expression and comprehension of communicative intentions (intention-reading)” (2003, p. 325). Moreover, he also claims his approach is construction-based. This, he surmises, has two important implications: “it means that the focus is on whole utterances and constructions – not isolated words or morphemes – as the most fundamental units of language acquisition [and] that we focus on children’s learning and use of particular words, phrases, and expressions, as concrete linguistic entities” (2003, p. 326). While Tomasello doesn’t breach the topic of SLA acquisition, if his point that language is theory of mind and pattern finding is adopted, a teaching curriculum with a fair amount of pragmatics, lexical grammar, and use of the kinds of vocabulary chunking found in corpus linguistics would be logical. In addition to this, the communicative approach would match his theory of intention-reading as it is connected with the abilities to interpret and learn the intentions of others underlain by communicative intentions.

By corpus linguistics it is implied that students study real, and authentic, examples of language use. Corpus-based linguistic study is done at the lexical, syntactic, and discourse level. In addition, Sinclair claims that meaning has an important effect on structure. If a word has two meanings, “it is possible to predict that it has two structures and this is only made possible by studying examples of language in use” (1997, p. 35). From this perspective, it is clear that teaching must be meaningful. It’s no use teaching an ESP writing class (for
example, to a group of travel agents) how to write academic papers. The class
time would be much better spent learning how to interpret and respond to faxes,
or how to write emails and doing tasks or activities that have real-life meaning for
the students.

**Connectionism**

Connectionist theories of the late 1980s and 90s attempted to narrow the
focus of the term *cognitivism*. Connectionism is a movement that attempts to
explain human intellectual abilities and the learning process using artificial neural
networks. Connectionists liken the brain to a computer, and artificial neural
networks are simplified models of the brain composed of large numbers of
nodes/units (the analogues of neurons) together with weights that measure
the strength of connections between the units. The links between units are
strengthened or weakened through activation or non-activation respectively.
Experiments on models of this kind have been used to demonstrate an ability to
acquire language. Connectionism differs from standard cognitive/nativist theories
in that it views learning on the basis of associated processes, rather than the
construction of abstract rules. The traditional view is that cognitive processes are
assumed to be in a serial order, whereas the connectionist model views cognitive
processes to be parallel. In relating connectionism to L1 acquisition Rumelhart
and McClelland (1986) developed a new computational framework that supposedly
mimicked neural networks for understanding the cognitive processes. Their model
reproduced closely the way in which children acquire the past tense in English.
The computer generalized on the basis of stored examples in a similar way to
children. In SLA as well, a number of researchers have explored connectionism.
In one study, Sokolik and Smith (1992) investigated the assignment of gender to
French nouns using a connectionist framework. They devised a computer-based
connectionist type network model that learned to identify correctly, the gender of
a set of French nouns. The model was then able to generalize from that learning
experience and assign gender to previously unstudied nouns with a higher degree
of reliability. Learning, in this view, is thought to take place as the strength of
given inter-connections between nodes increases as the associative patterns are
repeated over time. Key to this type of learning is that the learner does not extract
rules and then apply them, but merely registers associative patterns that strengthen
with use. The analogy of a computer’s cache memory comes to mind. When you type in a URL on a computer, it relays messages to different points on the Internet ahead magically displaying the web page of the assigned URL. A simplified example of a message-relaying path would include a request from your computer to your host provider, from your host provider to an international server, and from the international server to the server of the web page of the URL you typed. The return trip is made, and thus the page appears (in milliseconds) on your computer screen. Without the use of cache memory, the process of accessing web pages in your computer’s browser takes more time. However, the cache memory on your computer records all the information from all the servers required to access the websites you frequent most. By doing this it saves time when you connect to the same web page, recognizing the path of the URL before it is even completely typed in your browser. In making the analogy with connectionism, the various Internet servers your URL request goes through can represent the nodes, units, and neurons of your brain. The strength of connection between these servers (and therefore the speed of connection between your computer and the web page) is demonstrated by the frequency of visits from your computer to the URL (this is accounted for in the cache memory). The computer can communicate more fluently with the various Internet servers that are recorded in the computer’s hard drive (the web pages that are “activated” in the computer’s cache memory) than with servers that are not recorded (or web pages that are “non-activated” in the cache).

Before discussing the issue of implications of the connectionist model to TESL methods, it is important to discuss the ways in which SLA may differ from first language acquisition. The first consideration is that of the patterns themselves. L1 patterns may or may not transfer to L2 and vice versa. Moreover, the neurophysiological changes or cognitive developments not related specifically to language may limit the learner’s ability to acquire language or may predispose the learner to particular acquisition strategies. In addition, contextual factors, such as the acquisition setting or the communicative demands placed on the learner, may affect acquisition. These factors need to be taken into account when applying a connectionist model to TESL methods. When the connectionist model is applied, a number of implications arise. The first involves the definition of a linguistic item. In pre-Chomsky days, linguistic items were fairly easily defined
(phonemes, morphemes, structures). Generative grammar led to a major emphasis on the grammatical rules as critical items to be learned as part of the knowledge base. Connectionist approaches suggest an even more complex composition of phonetic and lexical features as part as part of the base of linguistic knowledge.

In response to Rumelhart and McClelland’s research, Pinker and Prince (1988) point out that the model does a poor job of generalizing to some novel regular verbs. They claim that networks may be good at making associations and matching patterns, but they have fundamental limitations in mastering general rules such as the formation of the regular past tense. Therefore one of the biggest strikes against connectionism, according to them, is that the nets are not good at the kind of ‘rule-based’ processing that is thought to be fundamental to language learning. Nets do not take into account other factors of the brain such as different types of neurons and the influence of hormones. Others argue that the analogy to neural networks in the brain is too loose. The connectionist model is not without its critics. However, it comes as no surprise that most of the critical waves created ultimately emanate from the Chomsky/neo-Chomsky nativist camp.

The implications of connectionist theory for SLA teaching models include a reliance on whole language and lexical patterns over traditional form-focused instruction. As the model is data driven, clearly linguistic input heard by the language learner is essential. The theory suggests that what goes in as linguistic input, come out as well-formed language. Connectionist models of teaching should be designed so that the lexical patterns are easy to memorize. Miller’s (1956) Magical Number Seven experiment indicates levels of information retention. Therefore, when teaching or reinforcing lexical patterns it is best to organize them with this principle in mind. When encouraging the reinforcement of lexical patterns, students need to be given the chance to use the language in meaningful contexts. This will encourage the weighting and activation of lexical patterns in the brain. The use of corpora in presenting lexical patterns is useful, in that the learner will be exposed to commonly used lexis.

**Discussion**

While Chomsky’s original ideas of innatism have arguably run their course, the baby has not been thrown out with the bath water. Where then, do we stand
today in the world of theories of language acquisition? What direction are we
going in? I still see only two conflicting theories: nature vs. nurture or, in this
case, nativist vs. non-nativist theories. On one side, we have the nativists led by
the likes of Chomsky and Pinker. On the other, we have the non-nativists, led by
Ellis (2003) and his various tribes of constructivists. The recent thrust in both
of these sides can be attributed to the use of cognitive science to strengthen their
positions. Pinker (1995) describes the advancements made on both sides and
attributes them to “a diverse set of conceptual and methodological tools [that] has
been used to trap the elusive answers to our questions: neurobiology, ethnology,
linguistic theory, naturalistic and experimental child psychology, cognitive
psychology, philosophy of induction, theoretical and applied computer science.
Language acquisition, then, is one of the best examples of the indispensability of
the multidisciplinary approach called cognitive science” (1995, p. 176). The future
of these theories as they stand today is unclear. However, it seems likely that
someday an accommodation will arise. Clearly, input is essential in first language
acquisition, but this input must somehow supply data for a linguistic mechanism
that has evolved in the human species and that is possessed by all humans. Fifty
years from now our grandchildren might know which of these theories is correct,
or perhaps there is a happy medium.

From the perspective of language acquisition as drawing upon multiple
disciplines it is worth considering the findings of physicists at the University
of Cambridge in the UK. Josephson (1997), describes his integrated theory of
nervous system functioning embracing nativism and constructivism, by comparing
algorithmic and constructivist approaches and then presenting his third option,
an integrated approach. In applying this approach to language, Josephson adapts
a proposal of Elman: “If children develop a robust drive to solve [the problem
of mapping non-linear thoughts on to a highly constrained linear channel],
and are born with processing tools to solve it, then the rest may simply follow
because it is the natural solution to that particular mapping process” (Elman et al.,
1996). Josephson’s adaptation of this says, “if the possibility of innate systems
dedicated to language is not ruled out in accord with the dictates of the evangelical
constructivists, then we can change the above account to allow for a collection of
specific drives relevant to aspects of language acquisition, and for specific tools
that take into account universals of linguistic structure, and the corresponding
classifications” (Josephson, 1997). Josephson concludes that a “dynamical systems approach can integrate the algorithmic and neural network approaches to development, permitting the respective advantages of both schemes to be both retained and integrated” (1997). The implications in practice for language have been made by Josephson and Blair (1996). They call for a holistic approach to language, and while their original research may be considered ahead of its time, they seem fully aware of this: “it is not feasible at this stage in the research to attempt to make the theory perfect and to study exhaustively all permutations and combinations of the concepts involved.... The present work can be only a first approximation to the truth, and future research will be needed to amend any errors and add further necessary detail” (Josephson & Blair 1996).

**Conclusion**

The evolutionary-nativists and the non-nativist constructivists have given the world of applied linguistics a breath of fresh air. While remnants of Chomsky’s ideas of innatism remain with the nativists, the thrust of momentum in constructivism gives linguists a chance to explore theories outside of the Chomsky box. The current constructivist thrust represents a swing in the pendulum away from Chomsky’s theoretical dominance. The expanse of the constructivist tribe overlaps into many disciplines. The future of applied linguistics is on the verge of yet another new frontier. One can hope that this new frontier doesn’t stagnate and linger for decades as Chomsky’s did (and arguably still does). The complexities of first and second language acquisition may or may not be fully understood in our lifetime. However, as linguists and teachers it is our responsibility to continue the search for answers, using all of the knowledge accrued and all of the technology available.

**References**


The Relevance of Chomsky in 21st Century Second Language Acquisition