MASSIMO’S CHANGES IN RED

TOM’S CHANGES IN TEAL

Draft BB commentary proposal on Christiansen and Chater

Here we go again: language is the result of something else: General cognitive operations, statistical generalizations, learning principles, working memory, chunking, cultural constructions. Such attempts have a long long history, e.g., Rousseau (xxx) and a considerable modern literature on the issue started in the 1970s. (e.g., Bever, 1970; Bates and MacWhinney, xxx; Rummelhart and MClelland, xxxx;Hawkins,xxxx; notable recent examples include Arbib, 2012; Bybee, 2007; Rizzolatti & Arbib, 1998; Tomasello, 2003, 2006). (Perfors, Tenenbaum, & Regier, 2011) (Reali & Christiansen, 2005) (Christiansen and Chater XXX). All these attempts have been quickly and persuasively countered (Wexler, 2002) (Kam & Fodor, 2013) (CRAIN, THORNTON, & KHLENTZOS, 2009; Gualmini & Crain, 2005) (Berwick, Chomsky, & Piattelli-Palmarini, 2013; Pietroski, 2008) Piattelli Palmarini & Bever, 2002; Piattelli-Palmarini, Hancock, & Bever, 2008).

I. So, here *we* have to go again. Monotonously, but efficiently, the fundamental counter arguments are: (1) Linguistic data from a variety of languages that cannot be explained in that way; (2) Specific brain activation patterns unique to language; (3) Pathological cases showing double dissociation between the loss of linguistic ability and loss of cognitive abilities.

1. A brief selection of apparently irreducible linguistic systems follows: in each case, there is no consistent general explanation to be found in a domain of cognition outside of language. And any attempt at specific cognitive explanations for each separate system would also require a more general theory of why a given cognitive explanation is invoked for one system and not for others.

(a) The verb-object constraint (VOC) (Baker, 2008, 2013). In our conceptualization of the world, actions are more intimately connected with their agent than with the object, but not syntactically so. Syntactically, verbs are more intimately connected with their object (the internal thematic role of the verb) than to the subject. Verb+Complement forms a syntactic constituent (a chunk) but Subject+verb does not. This explains the fact that in all languages of the world idioms are formed by a verb and its object (*kick the bucket, sell the farm, hits the fan etc.*)MASSIMO, COULD YOU ADD SOME EXAMPLES FROM OTHER LANGUAGES HERE: never by a subject and the verb.

(b) The universal ordering of the “top” of the sentence (the so-called left periphery). (Belletti, 2004; G. Cinque, 1999; G. Cinque, 2013) The universal hierarchy of syntactic and semantic dominance relations tells us that > indicates dominance in the ordering of modal modifications of sentence, a transitive relation), for instance: evidential (*allegedly*) > epistemic (*probably*) > necessity *(necessarily*) > continuative (*still*) > durative (*briefly) >* obligation (*obligatorily*) > completive *(partially*). For example, in English we have:

*Jim is allegedly probably unable to frequently deliver assignments on time.*

*\*Jim is frequently unable to probably deliver allegedly his assignments on time.*

There many examples from other languages (e.g., Italian, Bosnian/Croatian/Serbian, Hebrew, Chinese and more) in the 4 collective volumes on the structure of the left periphery, suggesting that this ordering is universal. Explanations based on statistical regularity, general cognition, pure logic or social conventions appear utterly implausible. an orderingsyntactic

(c) Scarcity of grammatical encoding

It is also important to consider the arguments of Cinque (2013), who points out that “only a fraction of our cognitive concepts and distinctions seems to find distinctive grammatical encoding in the languages of the world [...] in one of the closed classes of categories (affixes, particles, auxiliaries, prepositions, etc.) that belong to the functional rather than the substantive lexicon [...] Most cognitive concepts and distinctions do not find any such encoding. (Cinque 2013:1)”

“Verbal projections in clauses grammatically encode (through affixes, particles, auxiliaries, etc.) distinctions relating to the internal and external temporal consistency of events (tense and aspect) and the speaker’s attitude toward the truth of the proposition (mood), but they are never found to encode such human cognitive universals as “shame”, “mourning”, “sexual taboos”, etc., nor otherwise cognitively significant concepts like “worry” “peril”, “fear”, “hunger”, “love”, “death”, “awe of god”, etc.” (2013:3-4)

(d) Conceptually possible, but linguistically impossible word ordering.

An even more severe problem arises from Cinque’s observation of universal ordering constraints, which make little sense from the point of view of cultural evolution of a learned system: “many potential orders are never found [...] which poses a puzzle for any culturally based account.” (Cinque 2013: 17)

DAVE 2 OR EXAMPLES OF REAL AND NEVER FOUND ORDERINGS OF DET NUM ADJ NOUN. RENDERED IN ENGLISH

2. There are specific brain activation patterns for language, different from brain activations for other cognitive processes.

A cardinal example: Musso, Moro et al. (Musso et al., 2003) have shown activation of Broca’s region when subjects are checking the correct or incorrect applications of just learned real grammars, but inactivation of that area when they check for correct or incorrect applications of arbitrary (in fact linguistically impossible) grammars. Logarithmically increasing activation of specific brain areas has been ascertained for real sentences and jabberwocky sentences as contrasted with word lists and number sequences (Pallier, Devauchelle, & Dehaene, 2011; Perani et al., 1996).

3. There is a vast literature on various kinds of aphasias selectively bearing on specific kinds of syntactic forms (notably object relatives, object gaps and object clefts) and on selective losses of kinds of lexical elements (verbs but not nouns, nouns but not verbs, names of persons only, names of famous monuments only, names of Countries only and so on) (Semenza, Zettin, & Borgo, 1998) show the high specifity and modularity of language, at different levels (Curtiss, 2013). As to the mirror image of such deficits, when language is either intact (as in Williams Syndrome (Zukowski, 2005, (2003)) or even enhanced (Smith & Tsimpli, 1995, 1997; Smith, Tsimpli, & Ouhalla, 1993; Tsimpli & Smith, 1998) in subjects with impaired general cognitive abilities. (For general review of many more such double dissociations, see CurtissXXX, Van Dely….StromswoldXXX)

In brief, there are numerous language structural specific phenomena that resist external explanation, enough to show that language has a unique cause. This is not to say that no universal properties of attested languages may have some specific cognitive explanations, (See the next section for a discussion of some of these.)

II We now turn to specific critiques of the target article.

C&C present two main points, each with two steps: 1a) working memory constraints account for many features of sentence processing during comprehension; 1b) these features in turn can account for a variety of universal properties of language. 2a) Thus, learning a language is actually learning a set of rapidly deployable recoding templates and processes; 2b) and what appears to be “structure” of language is actually culturally and historically determined.

We will respond with two corresponding kinds of discussions. 1a) There is a long history of explorations of the role of working memory constraints on language processing, e.g., from George Miller, 1951 to (Ullman, 2001) and many other authors. There is a corresponding set of explorations of how the constraints are reflected in attested languages (Bever, 1970, 2013; Hawkins 1983, 1994, 2004). C&C’s summary of these considerations is good, albeit historically myopic. 2a) The conclusion that language learning is *just* pattern learning is a non sequitur; we review facts and logical arguments that children are learning a grammar as well as a set of processing procedures. 2b) The conclusion that language properties are *just* culturally accumulated is also a non sequitur, with considerable forms of evidence and logical arguments against it.

We find what they say quite reasonable on the role of cognitive and behavioral constraints until approximately page 19, where they insist a tight time bottleneck (1) forces language acquisition to be a matter of learning a skill, and (2) that learning can only happen in real time. (this is highly suspicious; infants and young children are observed to ‘practice’ their language, including babbling by congenitally deaf infants (Weir, XXX. Petitto & Marentette, 1991) Moreover, the highest-level chunks ought to remain more accessible than written-over, passed-on lower-level chunks.) SOMETHING IS MISSING HERE

The authors attempt to bridge the gap between psycholinguistic research and syntactic theory in terms of general memory constraints. Put in that perspective, their claim could be stated as an endorsement of *recency* as the central memory factor constraining fast processing: only the last few elements parsed can be relevant to local, low-level processing. This is an important observation, as far as it goes; see below for our discussion of the extensive literature already exploring these ideas.

However, their focus on the “here and now” (which, upon reflection, must mean “only the recent past”), and the implicit emphasis on “habits of processing”, is somewhat myopic. Recency and frequency are only two of several major factors determining accessibility to memory. Another is primacy: from a presented a string of unstructured elements, recall will hold accurately only of the first few, and last few elements. Indeed, evidence suggests that the less structured the input, the greater the relative weight of primacy over recency in determining its accessibility. Applying this to language, we then expect a special role for the “left edge” in early acquisition (when the input cannot yet be fully processed, and so appears maximally unstructured). Novelty is another key property of human memory, which can apparently circumvent the “here and now” bottleneck of linguistic memory. For example, encountering a speaker with an accent, one may clearly remember low-level details of pronunciation – because it is unusual. Finally, salience is another factor independently determining access to memory; for example, one may remember the exact words and tone of utterances during a highly emotional event. Plausibly, this dimension of memory access finds a linguistic expression in phenomena such as contrastive focus and other information structure effects.

Facts and logic indicate that it must be true that one part of language acquisition must indeed be to assign the right internal boundaries (“chunks”) to external stimuli. That chunking is an inextricable feature of language processing goes back to Miller, 1955 and others (e.g., Carroll and Bever, 1972; ????) In structural theories, such as Generative Grammar, notably in the recent Minimalist Program, cyclic points of closure (called phases) in the derivation of complex sentences are currently countenanced, with many subtle and experimentally validated consequences (Chomsky, 2001; Epstein & Seely, 2002) (Gallego, 2012). This approach is totally alien to the idea that the character of phases may be the result of statistical generalizations. One thing is to admit that some kind of “chunking” of the incoming linguistic stimuli may be an essential part of acquisition and production, but it is another thing to suppose that language acquisition is effectively *exhausted* by this process, or that the process is one of “learning” as such. Indeed, infants seem to acquire the prosodic contour of their native language even before birth, strongly constraining or even dictating the relevant “chunk” boundaries (and other nuanced details of the language) well before any substantive “item-based constructions” can be postulated. It is precisely in ignoring this kind of instinctual cueing of language structure that C&C’s account goes awry.

CUT ~~“our account is agnostic with respect to the specific characterization of the various levels of linguistic representation, e.g., whether sound-based chunks take the form of phonemes, syllables, etc. What is central for the Chunk-and-Pass account is that there should be some form of sound-based level of chunking (or visual-based in the case of sign language), and a sequence of increasingly abstract levels of chunked representations into which the input is continually recoded.” (p. 13)~~

~~Indeed, the account has little bite without something more being said about the nature of the chunks.~~

~~“Importantly, though, this does not imply that linguistic relations are restricted only to adjacent elements but, instead, that they may be formed between any of the small number of elements maintained at a given level of representation.”~~

While their account helps them escape some empirical problems, it highlights other, more severe ones. It says nothing, for instance, about which of the small number of elements at a given level may relate to each other; as far as C&C are concerned, anything goes. This flies in the face of considerable evidence for richly nuanced and apparently universal principles governing such relations (subjacency, case, theta relations, etc.). ~~; see below.~~ It also makes the ubiquitous presence of long-distance dependencies mysterious. If learners look first for local associations in blindly segmenting their language, subject to a crippling limit on short-term memory, it is unclear why or how long-distance extraction ought to be stable in any lineage, much less universal. This is especially true for so-called wh-in situ languages, where the long-distance dependency holds between a filled argument position and a possibly distant scope position (which may not even be explicitly marked); we simply see no way that a naïve learner would be led to induce such patterns from raw data without guidance of an internal Bauplan.

C&C assert “The ability to acquire and rapidly deploy a hierarchy of chunks at different linguistic scales is parallel to the ability to chunk sequences of motor movements, numbers or chess positions: it is a skill, built up by continual practice…. If linguistic input is available only fleetingly, then any learning must occur while that information is present; i.e., learning must occur in real-time, as the Chunk-and-Pass process takes place. That is, any modifications to the learner’s cognitive system in light of processing must, according to the Now-or-Never bottleneck, occur at the time of processing.” (18-19)

Consider the following parallel argument about visual processing: we know that visual information is available only fleetingly, rapidly overwritten by the flood of new information. This means that learning to process the visual scene must happen online, and that vision is a learned skill. In this context, it is clear that we are simply begging the question of whether innate machinery dictates how visual input is processed, or whether it “emerges” from learning to process the “buzzing, blooming confusion” of raw sensory data. In the case of vision, we know that *how the input is perceived* relies on strongly innate mechanisms – cell populations dedicated to detecting low-level features (REFS). It is an open question whether invariants of implementation in the brain pre-determine the range of possible structures in language.

They cite the remarkable case of SF, who learned to hierarchically chunk very long sequences of digits, as support for their general idea. But more thoughtful examination of the facts of SF’s case point in a different direction. First, SF was able to achieve his remarkable chunking only by *interpreting* the digits as meaningful units (running times, for example), arguably engaging at least some of the specialized machinery of language. Indeed, it is a recurring theme of research on memory that short term memory, held in Baddeley’s (1986) “phonological loop”, engages specialized language production mechanisms – in effect, sequences in a short-term memory task are maintained in memory by covert internal repetition). It is not clear, then, if this remarkable anecdote sheds any light on non-language-specific cognitive architecture.

~~“Whatever the appropriate computational framework, the Now-or-Never bottleneck requires that language acquisition should be viewed as a type of skill learning, such as learning to drive, juggle, play the violin, or to play chess.” (19)~~

~~‘The challenge of language acquisition is to learn a dazzling sequence of rapid processing operations, rather than conjecturing a correct “linguistic theory.”’(p. 20)   
  
Maybe, or maybe not, but we do not see how this follows from the “Now or Never bottleneck.” Indeed, it seems merely to beg the central question it claims to answer, namely, to~~

The central issue is to what extent are the chunks posited by the child bootstrapped from the data itself with no prior “theory of language”, and to what extent are they guided by innate instincts that reject data not fitting the prior knowledge of language those instincts instantiate? To make this clear, it is certainly conceivable in the abstract that the challenge of language acquisition is solved not by general-purpose skill-learning, but by deploying instinctual control mechanisms, akin to limb control or visual processing.

“Nonetheless, on-line learning contrasts with traditional approaches in which the structure of the language is learned “off-line” by the cognitive system acquiring a “corpus” of past linguistic inputs and choosing the grammar or other model of the language that best fits with those inputs….Crucially, such information is not, therefore, in a suitably “neutral” format to allow for the discovery of previously unsuspected linguistic regularities. In a nutshell: the “lossy” compression of the linguistic input is achieved by applying the learner’s current model of the language.”(20-21)

Indeed. We take this to be one of several strong arguments for the prior presence of specialized mechanisms attuned to the particular kind of regularities found in language. Put another way, they are saying the system cannot search at leisure for regularities that are not already available to it; this is at least consistent with the possibility that such regularities are not learned at all, but enforced by the very architecture of the system.

~~Discussing re-processing, they assert “The on-line learning constraint is that material is learned only when it is being processed—ruling out any putative learning processes that involve carrying out linguistic analyses or compiling statistics over a stored corpus of linguistic material.” (21)~~

~~We find this quite reasonable, and are puzzled by their quick turn to endorsing statistical methods as a way of extracting the relevant regularities~~

~~Discussion near top of page 21: Learning must be online; now or never bottleneck means learner can only apply their current model of language, losing things in compression that would point way to more correct model. So, no at-leisure examination of corpora for statistical analysis, new generalizations.  
This again seems to discount the very possibility they take it to argue against, namely, that significant features of the model of language are present in advance of encountering the input. That the system is keyed to recognize certain kinds of regularities (as specifically linguistic regularities), and indeed this may be the basis for inducing the categories the grammar is stated over. Grammars aren’t invented as intellectual objects, to cover generalizations we notice; rather, generalizations in the data cue us to infer certain grammars instinctually and inevitably.  
 More broadly, the “chunking” in their “chunk and pass” mantra is a more sophisticated process than they seem willing to admit. Indeed, one might better call it “analysis” than “chunking”; a critical question is the nature of this analysis, whether and how it is guided by internal principles antecedent and insensitive to environmental input.~~

~~p.22 the nod to exemplar dogma risks incoherence. The previous few pages argued that learning must occur online during processing (or during private rehearsal of processing). But on this page, we suddenly hear that “One way to learn in a local fashion is to store individual examples (this requires, in our framework, that those examples have been abstractly recoded by successive Chunk-and-Pass operations, of course); and then to generalize piecemeal from these examples.”  
So… offline lexicon-wide generalization is ok, but only when it’s your f(l)avor? I’m baffled. Surely more abstract representations are more easily accessible on anyone’s model.  
  
p. 22: as footnote 14 at least concedes, the (straw man) viewpoint challenged by this “new” insight corresponds to the Aspects-era view of child as a “mini-linguist”, a position to which few if any modern theorists cling. P&P style triggers do indeed look more compatible with the Now or Never Bottleneck (though the troubles with parameters as currently formulated are serious). By that token, little new is actually being contributed to the debate here, a complaint that applies more broadly.  
        The need to “chunk and pass” in light of severe short-term memory limitations and rapid input is hardly a new idea; we see it clearly already in the discussion motivating Frazier and Fodor’s (1978) parsing model, with a preliminary phrase packager (PPP) chunking the input and passing it on to their sentence structure supervisor (SSS).  
  
“By contrast, the principle of local learning is respected by other approaches. [… These] accounts of language acquisition tie learning and processing together—and assume that language is acquired piecemeal, in the absence of an underlying bauplan.” (p. 23)  
But that’s a feature of those accounts, not an argument as such that there is no Bauplan. One could perfectly well imagine local learning guided by an innate Bauplan; the issues are distinct.  
  
“local chunks can be defined at many different levels of abstraction […] and generalizations from past processing to present processing will operate across all of these levels.” (p. 23)  How is this not simply contradicting what they say about now or never learning?  
  
p 25: ah, here it comes, the triumphal crow of the language-is-culture enthusiasts. Lexical items reproduce and proliferate or die across generations of speakers. As grand a space to find explanation in as neo-Darwinian theory, and as deeply flawed for the same reason: it pays no attention to what the actual limits on variation are, assuming all form is the result of evolution acting on an unstructured “blank slate”. Deeply misguided: in biology generally, and likely also in language, evolution cannot be understood except in light of the limits of canalization.~~

~~p. 26 “Expanding on the notion of reduction and erosion, we suggest that constraints from Chunk-and-Pass processing can provide a cognitive foundation for grammaticalization (Hopper & Traugott, 1993). Specifically, chunks at each level of linguistic structure—discourse, syntax, morphology, and phonology—are potentially subject to reduction.”~~

~~Grammaticalization provides a particularly interesting testing ground for their ideas. As they say, their model predicts general reducibility for the most frequent forms. In this light,~~

~~“For example, computational phylogenetic analyses indicate that word order correlations are lineage-specific (Dunn, Greenhill, Levinson & Gray, 2011), shaped by particular histories of cultural evolution, rather than following universal patterns as would be expected if they were the result of innate linguistic constraints (e.g., Baker, 2001) or language-specific performance limitations (e.g., Hawkins, 2009).”(28) The study they cite is deeply flawed. Dunn et al examined correlations of pairwise orderings, which illegitimately groups together unlike word orders. See Cinque (2013) for discussion.~~

~~p. 28: they expect few if any true language universals, aligning themselves with Bybee and Evans & Levinson. Again, I think Universal 20, as typological generalization and (especially) as artificial language learning effect, is a clear counter-example.   
That’s the least of their problems. Especially severe is the considerable evidence for semantic uniformity; that seems totally incompatible with Bauplan-less, item-based learning.~~

~~“It remains, of course, of great interest to understand the biological evolutionary history that led to the cognitive pre-requisites for the cultural evolution of language. Candidate mechanisms include joint attention, large long-term memory, sequence processing ability, appropriate articulatory machinery, auditory processing systems, and so on. But this is the study not of language evolution, but of the evolution of the biological precursors of language (see Christiansen & Chater, 2008; and Pinker & Bloom, 1990, for an opposing perspective).”(28-29)~~

~~This is to confuse E-language and I-language. “language change and language evolution” – same theory – this is sophistry. And confusion over evolution.~~

~~“the Chunk-and-Pass account is in line with other approaches that assign processing limitations and complexity as primary constraints on long-distance dependencies, thus potentially providing explanations for linguistic phenomena, such as subjacency (e.g., Berwick & Weinberg, 1984; Kluender & Kutas, 1993), island constraints (e.g., Hofmeister & Sag, 2010), referential binding (~~e.g., Culicover, 2013), ~~and scope effects (e.g., O’Grady, 2013).”(30-31)  
  
“Thus, patterns of regularity in language will emerge locally and bottom-up, from generalizations across individual instances, through processes of language use, acquisition, and change.”(33)   ...Assuming, as has not motivated, that there is no “Bauplan”, no innate guidance. As far as we can tell their approach simply begs one of the central questions of linguistic inquiry.~~

~~This evidence includes sharp-edged word order universals of the sort they (following Evans and Levinson 2009) dispute, and evidence of striking uniformity at the level of interpretation. Some kind of innate program or Bauplan is implicated, as a crucial support guiding “learning to process”. This conclusion~~

Our position is strengthened by suggestions from artificial language learning that monolingual adult speakers retain access to whatever biases determine the typological distribution of word orders, as argued by Culbertson and colleagues (Culbertson, Smolensky, and Legendre; Culbertson and Smolensky; Culbertson and Adger REF). This speaks directly against the idea that languages are collections of idiosyncratic cultural atoms shaped only by gradual and non-specific evolution; adults seem to “know” things about languages they have not encountered. This does not mean that processing is not ultimately at the heart of the story, but there must then be something more to say about processing, in terms of specifically linguistic architecture, to account for semantic uniformity and word order universals.

~~“For example, the frequent processing of simple transitive sentences, processed individually as multi-word chunks, such as ‘I want milk’, ‘I want candy’, might first lead to a wider track involving the item-based template ‘I want X’. Repeated use of this template along with others (e.g., ‘I like X’, ‘I see X’) might eventually give rise to a more abstract transitive generalization along the lines of N V N (a highway in terms of our track analogy). Similar proposals for the emergence of basic word order patterns have been proposed both within emergentist (e.g., O’Grady, 2005, 2013; Tomasello, 2003) and generative perspectives (e.g., Townsend & Bever, 2001). Importantly, however, just as with generalizations in perception and motor skills, the grammatical abstractions are not explicitly represented but result from the merging of item-based procedures for chunking. Thus, there is no representation of grammatical structure separate from processing. Learning to process is learning the grammar.” (35)~~

~~Again, this assumes that “analogy” and learning are sufficient explanations, and that what is learned is language and item-particular. This faces problems in light of the apparent sensitivity of adults to grammatical patterns in languages they have not encountered (for a classic demonstration, Musso, Moro et al 2003; also Culbertson et al), as well as in the deviations from adult patterns found in child language acquisition.~~

~~Re-construction is criticial. Not imitation, but creation anew, using similar cognitive resources. This is especially so because half of the picture is hidden; the ‘message’ is covert.~~

~~“To cope with the Now-or-Never bottleneck, the language system chunks new material as rapidly as possible at a range of increasingly abstract levels of representation. As a consequence, Chunk-and-Pass processing induces a multilevel structure over linguistic input. The history of the process of chunk-building can be viewed as analogous to a shallow surface structure in linguistics, and the repertoire of possible chunking mechanisms and the principles by which they can be combined can be viewed as defining a grammar. Indeed, we have suggested that chunking procedures may be one interpretation of the constructions that are at the core of linguistic theories of construction grammar.” (36)~~

~~This passage highlights the emptiness of the current proposal.~~ The authors seem impressed that by calling limitations on short-term working memory “the Now or Never bottleneck”, and parsing “Chunk and Pass”, they have uncovered something new and important. “chunking procedures may be one interpretation of the constructions that are at the core of linguistic theories of construction grammar.”

Parsing and constituent structure relate, and the boundaries the parser finds match up to boundaries between linguistic units; that is hardly new. One can link this to construction grammar (for counters to this approach see (Crain, 1991; CRAIN et al., 2009)) (note, though, that constructions would seem to correspond to “chunks”, not general chunking procedures), if one finds that fashionable. Or one can link it to many other kinds of grammatical architectures (e.g., Fodor and Bever, 1965; Bever Kirk and Lackner, 1972). But this focus on extracting regularities from the input draws attention away from the central issue, namely, whether and how biologically determined invariants of internal processes delimit the range and nature of languages that can be acquired.

~~“The history of the process of chunk-building can be viewed as analogous to a shallow surface structure in linguistics, and the repertoire of possible chunking mechanisms and the principles by which they can be combined can be viewed as defining a grammar.”~~

There is an important point, however, ~~here~~ that we largely agree with, namely, that processing is about grammar. However, there is something deeply wrong about asserting that what is at stake is just procedures of segmentation, inducing regularities in surface structure. Crucially, that is only one half of the picture, the “easy”, external portion of language. The other half of language is the interesting and “hard” bit, the internal computation of inner forms and meaning. Interpretation is not just a higher-level segmentation of the external form, and its essential characteristics must follow in large part from internal principles.

~~As one example of what we have in mind, consider the hierarchy of grammatical positions identified by Cinque (1999) and subsequent work. This exquisitely articulated, apparently universal structure is found to accurately describe language after language.~~

~~As another example, consider Keenan and Comrie’s (1977) Accessibility Hierarchy:~~

~~Subj > DirObj > IndObj > Oblique~~

~~Across languages, this hierarchy determines whether and how an NP bearing the indicated grammatical function may enter into A-bar relations. If wh-extraction can affect one position in the hierarchy, it may affect all higher ranked positions to its left. If a position may host a resumptive pronoun, so may any position to its right (Hawkins 2010).~~

~~Of course, this hierarchy could well be explained in terms of processing, with dependencies into lower-ranked positions placing greater demands on memory or being subject to greater interference. But such a processing account presupposes a theory that explains what a subject is, and how it is different from a direct object, and so on. Crucially, there must also be an explanation for their universal structuring, mapping these concepts to “chunks” in such a way that the subject is always more accessible for A-bar dependencies (which themselves must be discriminated properly from surface-learnable but unattested patterns), the object (but not the subject) is always available to determine idiomatic meanings in combination with the verb, and so on.~~

Arbib, M. A. (2012). *How the Bran Got Language: The Mirror System Hypothesis*. Oxford UK and New York: Oxford University Press.

Baker, M. C. (2008). *The Syntax of Agreement and Concord*. Cambridge UK: Cambridge University Press.

Baker, M. C. (2013). On agreement and its relationship to case: Some generative ideas and results. *Lingua, 130*(June ), 14-32.

Belletti, A. (Ed.). (2004). *Structures and Beyond. The Cartography of Syntactic Structures, vol. 3.* . New York: Oxford University Press.

Berwick, R. C., Chomsky, N., & Piattelli-Palmarini, M. (2013). Poverty of the Stimulus Stands: Why Recent Challenges Fail In M. Piattelli-Palmarini & R. C. Berwick (Eds.), *Rich Languages from Poor Inputs* (pp. 19-42). Oxford, UK: Oxford University Press.

Bybee, J. (2007). *Frequency of Use and the Organization of Language*. Oxford and New York: Oxford University Press.

Chomsky, N. (2001). Derivation by phase. In M. Kenstowicz (Ed.), *Ken Hale: A Life in Language* (pp. 1-52). Cambridge, MA: The M.I.T. Press.

Cinque, G. (1999). *Adverbs and Functional Heads a Cross-linguistic Perspective*: Oxford University Press.

Cinque, G. (2013). Cognition, universal grammar, and typological generalizations. *Lingua, 130*, 50-65. doi: 10.1016/j.lingua.2012.10.007

Crain, S. (1991). Language acquisition in the absence of experience. *Behavioral and Brain Sciences, 14*, 597-650.

CRAIN, S., THORNTON, R., & KHLENTZOS, D. (2009). The case of the missing generalizations. *Cognitive Linguistics, 20*(1), 145-155.

Curtiss, S. (2013). Revisiting modularity: Using language as window to the mind. In M. Piattelli-Palmarini & R. C. Berwick (Eds.), *Rich Languages from Poor Inputs* (pp. 68-90). Oxford, UK: Oxford University Press.

Epstein, S. D., & Seely, T. D. (Eds.). (2002). *Derivation and Explanation in the Minimalist Program*. Oxford UK: Blackwell.

Gallego, A. J. (Ed.). (2012). *Phases: Developing the Framework*. Berlin and New York: De Gruyter Mouton.

Gualmini, A., & Crain, S. (2005). The structure of children's linguistic knowledge. *Linguistic Inquiry, 36*(3), 463-474.

Kam, X.-N. C., & Fodor, J. D. (2013). Children's acquisition of syntax: Simple models are too simple. In M. Piattelli-Palmarini & R. C. Berwick (Eds.), *Rich Languages from Poor Inputs* (pp. 43-60). Oxford UK and New York: Oxford University Press.

Musso, M., Moro, A., Glauche, V., Rijntjes, M., Reichenbach, J., Buechel, C., & Weiller, C. (2003). Broca's area and the language instinct. *Nature Neuroscience, 6*(7), 774-781.

Pallier, C., Devauchelle, A.-D., & Dehaene, S. (2011). Cortical representation of the constituent structure of sentences. *Proceedings of the National Academy of Sciences, 108*(6), 2522-2527.

Perani, D., Dehaene, S., Grassi, F., Cohen, L., Cappa, S. F., Dupoux, E., . . . Mehler, J. (1996). Brain processing of native and foreign languages. *NeuroReports, 7*, 2439-2444.

Perfors, A., Tenenbaum, J. B., & Regier, T. (2011). The learnability of abstract syntactic principles. *Cognition, 118*(3), 306-338. doi: 10.1016/j.cognition.2010.11.001

Petitto, L. A., & Marentette, P. (1991). Babbling in the manual mode: Evidence for the ontogeny of language. *Science, 251*, 1483-1496.

Piattelli Palmarini, M., & Bever, T. G. (2002). The fractionation of miracles: Peer commentary of the article by Michael Arbib: “From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics”. *Behavioral and Brain Sciences.* Electronic supplements, from <http://www.bbsonline.org/Preprints/Arbib-05012002/Supplemental/Piattelli-Palmarini.html>

Piattelli-Palmarini, M., Hancock, R., & Bever, T. (2008). Language as ergonomic perfection (commentary of Christiansen and Chater “Language as shaped by the brain”). *Behavioral and Brain Sciences, 31*(05), 530-531.

Pietroski, P. (2008). Minimalist Meaning, Internalist Interpretation. *Biolinguistics, 2*(4), 317-341.

Reali, F., & Christiansen, M. H. (2005). Uncovering the richness of the stimulus: structure dependence and indirect statistical evidence. *Cognitive Science, 29*, 1007-1028.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp [Viewpoint]. *Trends in NeuroSciences, 21*(5), 188-194.

Semenza, C., Zettin, M., & Borgo, F. (1998). Names and identification: An access problem. *Neurocase, 4*(1), 45-53. doi: 10.1080/13554799808410606

Smith, N. V., & Tsimpli, I.-M. (1995). *The Mind of a Savant*. New York and Oxford, UK: Blackwell.

Smith, N. V., & Tsimpli, I.-M. (1997). Reply to Bates. *International Journal of Bilingualism*.

Smith, N. V., Tsimpli, I.-M., & Ouhalla. (1993). Learning the impossible: The acquisition of possible and impossible languages by a poliglot savant. *Lingua, 91*, 279-347.

Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.

Tomasello, M. (2006). Acquiring Linguistic Constructions. In W. Damon, R. Lerner, D. Kuhn & R. Siegler (Eds.), *Handbook of Child Psychology. 2. Cognition, Perception, and Language* (pp. 255-298). New York: Wiley.

Tsimpli, I.-M., & Smith, N. V. (1998). Modules and quasi-modules: language and the theory of mind in a polyglot savant. *Learning and Individual Differences, 10*, 193-215.

Ullman, M. T. (2001). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research, 30*, 37-69.

Wexler, K. (2002). Lenneberg's dream: Learning, normal language development and specific language impairment. In J. Schaffer & Y. Levy (Eds.), *Language Competence across Populations: Towards a Definition of Specific Language Impairment* (pp. 11-60). Mahwah, NJ: Erlbaum.

Zukowski, A. (2005). Knowledge of constraints on compounding in children and adolescents with Williams syndrome. *Journal of Speech, Language and Hearing Research, 48*(1), 79-83.

Zukowski, A. ((2003)). Investigating knowledge of complex syntax: Insights from experimental studies of Williams Syndrome. In M. Rice & S. Warren (Eds.), *Developmental Language Disorders: From Phenotypes to Etiologies*. Mahwah, N.J.: Lawrence Erlbaum.