★ FORUM ★

A Plea for *Why Only Us* (Berwick & Chomsky 2016)

Wolfgang Sperlich

In *Why Only Us* (Berwick & Chomsky 2016), Bob Berwick and Noam Chomsky (henceforth, B&C), as masters of the metaphor, hammer home three linguistic home truths: (1) language is hierarchical (not linear), (2) all we need is Merge, and (3) speech (and communication) is external to language. When this is set in the context of biological evolution, B&C do admit to some limitations like "biology is more like case law, not Newtonian physics" (p. 36). As such, they make a very good case that can stand up in any court of law, excepting of course in kangaroo courts that proliferate these days, online and offline. I shall address a few below but before I do, allow me a few general observations.

Attacking Noam Chomsky has long been an industry that is encouraged and sustained by political reactionaries *cum* pseudo-scientists, pretentiously disguised as scientific debate. If one could only discredit him as a scientist, he would suffer as a political animal —only neo-fascists will attack blindly, as they always have done. So let us find a few linguists *cum* biologists who can prove that Chomsky's views on language evolution are a load of rubbish. *QED*. Chomsky at a ripe old age is, of course, becoming vulnerable to attacks by younger wannabes snapping at his heels. One simply has to disagree with at least some points that B&C raise and one has established oneself has a potential successor, paying heed to the Popperian obsession that science is all about falsifying existing theories. Like there is a science fringe that is always attempting to falsify Einstein, however benign he was politically. Generally, however, sensible scientists do attempt to prove Einstein's theories—witness the recent discovery and measurement of gravitational waves that were predicted by Einstein. Now, I am not saying that Chomsky is on par with Einstein, but wouldn't it be nice if linguists, biologists, neuroscientists, and what have you would concentrate on proving Chomsky right—"for Chomsky is an honorable man", as a modern Shakespeare would have to say.

As such, it is quite sad to see that even formerly avowed collaborators with Chomsky, like the now somewhat damaged Marc Hauser, must write articles that disagree with a number of important points that B&C raise. I am not for one moment suggesting that B&C are beyond criticism—they are not, but let's not try to falsify their theories which they are the first ones to admit, are tinged with speculation that is inherent to the topic of language evolution. To go on the attack and discredit a speculative theory by pretending to advance facts of the matter – where there are none—is like taking the wind out of the opponents sails by shadowing the sailboat with an ocean liner (a metaphor Chomsky the sailor might appreciate). Let's keep the playing field level, as the English would say.

ISSN 1450–3417
Biolinguistics 10: 051–059, 2016
http://www.biolinguistics.eu

Chomsky is a good sport and enjoys a good scrap, and he can give as much as he takes. When B&C go on the counterattack, they do so with good humor, on occasion pointing out some of the more bizarre critiques they have to endure. Witness Tomasello's 'UG is dead' moniker, which B&C counter as saying that:

> If so, then there is of course no topic of the evolution of UG—that is, of the evolution of language in the only coherent sense. Rather, the emergence of language reduces to the evolution of cognitive processes—which cannot be seriously investigated for the reasons that Lewontin has explained.
>
> (Berwick & Chomsky 2016: 97)

I will return to the obsession by cognitivists to subsume language as a mere phenomenon of the cognitive apparatus, whatever that may be. Let us briefly consider another terrible sin perpetuated by B&C, namely that language has not evolved from communication. Vyvyan Evans reviews *WOU* and comes to the following conclusions (amongst others):

- "It's quite a stretch to suggest that language didn't evolve to enable communication."
- "Indeed, the book attempts to make a virtue of disagreeing with almost everyone on how language evolved."
- "The reader is asked to swallow the following unlikely implication of their logic: [L]anguage didn't evolve for communication, but rather for internal thought."

(Evans 2016, online)

These very blunt instruments used for critiquing a perfectly sensible theory should not be in the armory of sensible academics—or shall we call Evans an intellectual? Evans might be enamored by the proposal, but then let's shoot him down with Chomsky's dictum of many an intellectual being a 'commissar' to uphold the reactionary paradigm.

Next in line is the very curious case of Elliot Murphy, who on his blog starts his review as follows:

> Bob Berwick and Noam Chomsky's new book Why Only Us: Language and Evolution has been making the rounds. I assumed this book would just be a re-hash of the fairly tiresome, hyper-sceptical 'mystery of language evolution' perspective the authors usually adopt. And it is in some respects. But it also includes a surprisingly decent discussion of recent literature on animal cognition. Berwick's ideas come through more clearly throughout the text, typically backed up with the usual selection of Chomsky's rhetoric, Martian analogies, irony and so forth. But both authors only brush over their core question of how hierarchy is actually established, pointing languidly to 'some algorithm' responsible for labeling (p. 10). It should be stressed I think that even Chomsky's more recent technical work doesn't go far beyond this 'some algorithm' attitude (2013, 2015). From the perspective of brain dynamics, 'some algorithm' becomes capable of being explored in a number of interesting ways, as I mention here and here and in upcoming papers (see also Boeckx and Theofanopo[u]lou's useful response to the latter paper).
>
> (Murphy 2016a, online)

Wow! We are so grateful that the book in question slightly rises above the "fairly tiresome, hyper-sceptical 'mystery of language evolution' perspective the authors usually adopt". The rest of the article is merely a vehicle to launch his own theory which is interesting in some respects but even more tiresome in its relentless self-assertion and name dropping (note Boeckx being mentioned early on as the current guru of biolinguistics). One reader of Murphy's blog review is even emboldened to praise his piece as "[a]t last a non genuflected description of a Chomsky's book!".

What comes next is an even greater surprise, namely that Murphy is elevated to double authorship in the current volume of *Biolinguistics* (Murphy 2016b, 2016c), introducing one of his articles thus:

> My intention in this piece is to briefly outline a novel hypothesis regarding the neurobiological implementation of feature-set binding, the labeling of feature-sets, and the resolution of linguistic dependencies arising from the cyclic combination of these labeled objects. One of the numerous motivations for this was reading Robert C. Berwick & Noam Chomsky's (B&C) recent book *Why Only Us: Language and Evolution* (Berwick & Chomsky 2016; henceforth *WOU*), which struck me as moderately comprehensive in its interdisciplinary scope (including good critical commentary on recent work in comparative neuroprimatology and theoretical biology) but severely impoverished in its range of linking hypotheses between these disciplines.
>
> (Murphy 2016b: 6)

We are ever so grateful that he has slightly revised his opinion to "moderately comprehensive in its interdisciplinary scope"—maybe in an attempt to get his articles accepted by *Biolinguistics*, which has always acknowledged the debt (intellectual and scientific) of Noam Chomsky as the founding father of biolinguistics. I have commented before on this paradox inasmuch as *Biolinguistics* does publish true-blue biolinguistics papers, but on occasion gets it horribly wrong with articles that are anti-biolinguistics as much as anti-Chomsky. So let us hit back *ad hominem*, one more time, against one Edmund Blair Bolles, who writes a blog about 'the origins of speech' and reviews B&C as follows:

> I love the fact that in the beginning, and before there was any language, and in some "completely unknown way" we got the computational atoms that Merge assembles. So we start with a miracle. Words get their meanings by invoking these concepts. Thus, when I speak of the Hudson, or the Seine, or the Nile I am getting my meaning, not by pointing to a specific geographical entity, but by invoking an innate concept of river that is older than language. This kind of raw Platonism has appealed to many thinkers over the centuries, but I confess to always being a bit repelled by the sterility of the realm of forms. (Bolles 2016, online)

No wonder Mr. Bolles has some difficulty in believing a word of B&C, if he is concerned with 'the origins of speech'. I suppose one must forgive the 'speech = language' enthusiasts for their simplistic reasoning, just like the flat-earthlings could not believe that the world is round. Actually, I am not sure why B&C traverse this issue in such detail, that is, making the point that sound production and voice—and its evolutionary sequences—are common to many species,

notably, of course, to songbirds, but no species apart from humans have developed anything like language. Songbirds sing beautifully, whales and dolphins create under-water symphonies, and our cat meows in a way that annoys my wife, just like the harsh voice of certain humans can be extremely off-putting (listen to Hitler's speeches and wonder how such a terrible voice could enthuse millions of Germans, lest they were hypnotized). The point I am trying to make is that common sense absolutely forbids any connection between producing sounds (including human speech sounds) and the potential for language. B&C argue convincingly that the human propensity for voice modulation was a pre-existing tool that was later used for externalizing language as an erstwhile mental product. Proverbial bird-brains, as those detractors named above, seem to waste a lot of valuable time of B&C, who somehow feel compelled to prove them wrong.

So let us return to one who at least believes in some of the 'basic properties' of language as advanced by B&C, namely Marc Hauser, whose review is also published in *Biolinguistics*. He starts with a rare compliment, namely that *WOU* "is a wonderful, slim, engaging, and clearly written book" (Hauser 2016: 1). He then goes on to claim that *WOU* is based on the following five premises:

> (1) Merge is the essence of language.
> (2) No other animal has Merge.
> (3) No other hominid has Merge.
> (4) Due to the simplicity of Merge, it could evolve quickly, perhaps due to mutation.
> (5) Because you either have or don't have Merge (there is no demi-Merge), there is no option for proto-language.
>
> (Hauser 2016: 1)

Of these, Hauser can only accept 2 and 3. So what's wrong with 1? He doesn't really make a case apart from saying that there must be much more to language than Merge. Nevertheless, as Hauser does agree that Merge exists, what's wrong with 4 and 5? Here his main argument seems to be that since B&C also maintain that Merge must interface with CI and SM, Merge cannot emerge (so to speak) by itself without parallel evolution of CI and SM. Given that B&C move ever closer to the idea that 'language of thought' equals CI, Hauser baulks at the idea, saying that "Language of Thought implies that the system is explicitly linguistic, and I don't believe it is" (p. 4). As I argue (and have argued before) that language = CI, one can, of course, dismiss Hauser forthwith. Hauser is also a fan (as he has to be, as an animal cognitivist) that SM is intimately connected with language (and Merge), hence we cannot dismiss the evolution of SM as paralleling language. Since B&C make a case for dismissing at least the externalization effect of SM as being related to the language faculty, Hauser's songbird obsession shines through, making the startling claim that "in particular, songbirds learn their song in some of the same ways as young children learn language" (p. 2; see also my comments on 'learning' below).

Having surveyed a few reviews, allow me to now write my own, with my first admission being that I agree with everything that B&C have to say, with the exception of various sections I do not really understand due to lack of technical knowledge. For example, the section on computer modeling of language and

cognition: I was naïve enough to assume that brain and language computations will perhaps never be replicated by a machine, but here B&C surprise me with "the well-known challenge is that there are many, many algorithms and implementations that can do the job" (p. 132). Nor am I *au-fait* with the current neurophysiology of the brain in humans and other species (such as songbirds). My real interest is in advancing B&C two steps further, namely, first, in equating human cognition with language (language equals thought) and, second, in pushing the anti-lexicalist ideas that minimize the problem of the 'lexicon'.

So what is my argument for the first assertion? The Cartesian proposition of *cogito, ergo sum* may be the best evidence for equating language with thought, for how else could you express this idea, if not by and through language. No language, no thought. Nobody has ever isolated a thought without language. Let us restrict the meaning of cognition to the ability to think. Learning to perform tasks without thinking is as such outside the realm of cognition. Practically all species are capable in some way of such learning but only humans can ask afterwards, "Now what do you think, did I learn it well?". Hence the proverbial 'teaching the monkey to perform tricks', or as B&C put it:

> If we reflect on this for moment, it appears that chimpanzees are perfect examples of pure 'associationist learners'—what they seem to have are direct connections between particular external stimuli and their signs. They do not seem to regard the apple they see in some mind-dependent way, as discussed in chapter 3. Rather, they have stored a list of explicit, mind-independent associations between objects in the external world and the ASL signs for them. This is far from human-like language ability—the chimps lack both Merge and the word-like elements that people have. If so, chimps are also eliminated as suspects in our whodunit.
>
> (Berwick & Chomsky 2016: 146)

Associated with this language = thought is, of course, the unpalatable consequence that language is not primarily a tool for communication. The famous witticism promulgated by Bronowski (1977, cited in Fujita 2009) is that were it true, then the first human uttering a word or phrase would not have a counterpart to understand anything (i.e. we have the first communication breakdown in human history). It makes sense to assume that the development of Merge established mental concepts that equated to language, allowing for initially simple abstractions leading to propositional thoughts. That such a development in a few individuals led to a selective advantage would equally make sense. The idea of externalizing such propositions to check if fellow individuals might have the same or similar thoughts would be a next step but fraught with many obstacles. To externalize mental language (= thought) into speech would have to be met with many frustrations along the way, like the communication breakdowns alluded to above. Exchanging thoughts via speech no doubt creates new feedback systems that give rise to new and possibly more interesting thoughts. A negative corollary might have been that such communication could be used for nefarious purposes. Animals are not known to communicate false warning signals but humans are. Communication, as a worst case scenario, developed into a narcissistic enterprise that is evident today as much as it has been throughout recorded history. The great communicators of our day, from Hitler to Reagan, used

speech-making as some sort of hypnotic mass medium, communicating precisely nothing but themselves. These people, as the proverb goes quite succinctly, do not think before they speak. They are like trained parrots who drill holes into the brains of their adoring fans, a feat otherwise known as brainwashing. Sure there is also the opposite effect; for example, communicating genuine feelings by saying so: "I love you!"

Let us also be clear what communication is not: the externalization of thoughts as self-reference, as typically achieved in the fields of science and literature. To externalize one's thoughts on how the universe works and how language might have evolved is to put on public record one's thoughts. Sure, scientists may talk to each other about research but the ultimate output is not to communicate to others what they found out—the output is a public statement of their thoughts. When B&C wrote *WOU*, they did not do so from a burning desire to communicate, they simply wanted to state the facts of the matter the way they see (=think) it. That their thoughts resonate with mine is not a matter of communication. I do not write this review in order to communicate with either A, B, or C. Externalizing one's thoughts in this way seems to be a good way to check the validity of one's thoughts for one has to translate one's language of thought into the product of writing. Writers who depend on their writing as a means to make a living will, of course, try to ensure that they have a wide readership—not to communicate with the readers but to entice them to part with their money to buy the book. Schrödinger famously dreamt his groundbreaking formula, and so do many other scientists and writers in terms of thinking for themselves—not to communicate with someone else. This whole issue about communication also harks back to the longstanding distinction between *langue* and *parole*, i.e. the latter being the use of language which in itself may be an interesting field of study but should not be confused with the study of language itself. As with Wittgenstein's game theory of what constitutes language, we all know the rules of the game but some players are better (or more devious) than others.

All in all, one cannot but vigorously defend B&C's following assertion:

> Accordingly, any approach to the 'evolution of language' that focuses on communication, or the sensorimotor system, or statistical properties of spoken language and the like, may be seriously misguided.
>
> (Berwick & Chomsky 2016: 84)

So what of the mysterious lexicon? B&C note:

> We will (speculatively) posit that the word-like elements, or at least their features as used by Merge, are somehow stored in the middle temporal cortex as the 'lexicon'—though as we mentioned in chapter 1, it is not clear how anything in memory is stored or retrieved.
>
> (Berwick & Chomsky 2016: 159)

This idea is reminiscent of the old idea that vocabulary items are endowed with mini-grammars, now called features or edge-features as used by Merge. The lexicon has always been a weak point in Chomskyan theory, so what about the simple (hence elegantly minimalist) solution to posit that there is no lexicon? This proposal seems to fly in the face of popular views on language, not to speak of

the publishing industry that makes a good living out of selling all manner of dictionaries. Nevertheless, the so-called anti-lexicalist stance has made headway in recent discussions, including by the above much maligned Murphy, who takes a somewhat hesitant step towards the proposition:

> These observations support Borer's (2005) exoskeletal morphological model which views open-class words as hidden 'conceptual packages' that are purely embedded in the syntactic structure, causing no alteration to it or itself. Only when the structure is built by phase is the package 'opened' (interpreted). This is one of many reasons why syntax appears to be entirely free of lexical influence, operating independently of the needs of feature matrices (see Epstein et al. 2014 on 'free' Simplest Merge). […] [W]ords are not concepts but rather *instructions to build concepts (from their semantic features)* [references excluded]. (Murphy 2016c: 30–31)

So, we still have 'words' but they are stripped of all syntactic edge features, reduced to 'conceptual packages'. How do they arise and where are they stored, if stored at all? Enter Fujita (2009) who, as far as I know, makes the strongest anti-lexicalist claim to date:

> *Basic claims of anti-lexicalism*
>
> a. Words are generated by syntax.
> b. The lexicon can be decomposed into FLN (Merge) and FLB (sound and meaning).
> c. Consequently, there is virtually no lexicon.
>
> (Fujita 2009: 143)

How exactly words can be generated by syntax remains a bit of a mystery:

> To a certain degree, it can be said that syntactic structure building by recursive Merge is at the same time a parallel hierarchical conceptual structure formation by Merging semantic atoms successively (say, conceptual Merge). This proposal, by no means, is intended to suggest that syntactic structure and semantic structure are the same, as was once claimed falsely by Generative Semantics. On the contrary, full semantic interpretation requires much more information than syntactic structure provides (in particular where the compositionality principle fails to capture the vastly multifaceted and flexible syntax-semantics relations), and syntax and semantics remain two autonomous modules as before. (Fujita 2009: 145)

So, what are these 'semantic atoms'? And why is it wrong to claim that "syntactic structure and semantic structure are the same"? Sure, Generative Semantics simply put the onus on semantics, but if you consider the proposition of language = thought, then why not claim that syntax = semantics? Fujita gives an example:

> Notice finally that to the extent that simple words are syntactically complex objects, it follows that Sub-Merge (Subassembly-type Merge) is always involved even in the derivation of two word utterances. This is so since to Merge *milk* and *cup* to form *milk cup*, for example, each of the two nouns must first be formed by Merge. (Fujita 2009: 148)

Does this mean now that this 'Sub-Merge' (no pun intended?) equals semantics? Fujita, in the end, cannot do any better than Murphy (2016c), who evokes 'conceptual packages' (see above):

> Needless to say, there has to be a universal pool of features in the human brain, different combinations of which will ultimately yield a different set of lexical items or words (sound–meaning pairings) available in particular I-languages. These are a residue of the lexicon that may safely be assumed to be part of FLB. (Fujita 2009: 143, fn. 11)

Are the sound–meaning pairings arbitrary? Why do the English say 'tree' and the Germans 'Baum'? Was there a proto-sound–meaning pairing? Let's say I am forming a thought in my head to the effect that as an externalization it reads 'This tree is beautiful'. Let's assume that this thought in my brain has no lexical equivalents. Only when I externalize this thought do I need lexical items which may well be inventions restricted only by the features of the system of externalization (some remnants like onomatopoeic words may well point to earlier vocalization features). Since externalization requires a raft of complex motor skills, one may then posit that such motor skills become part of the memory yielding a 'learnt' lexicon of a particular language. Even so, this learnt lexicon component is totally subject to syntax as proposed by Fujita and Murphy, one learns 'words' only in the context of syntax. It is well known that in the rare cases where children have no or only very limited lexical input for learning 'words', children will invent their own as required. I do realize that there is certain amount of circularity in this argument, since I cannot really make a case that this argument pre-existed as a mental construct totally devoid of lexical items. When I think, I do use ephemeral words—but note that babies born have language capacity but no lexicon in the sense used now. Since the language of thought may have a wider scope than the externalized language we use as active or passive comprehension, such a wider scope cannot, however, transcend the actual language we use (learnt lexicon included) every day, be it as linguists or tinkers, tailors, and candlestick makers for equal measure.

So, what could thoughts generated by Merge possibly be made of—biologically speaking? Murphy invokes 'brain oscillations' which still sounds like Newtonian physics to me (as alluded to by B&C before) and so I am somewhat surprised that none of the protagonists reviewed here (B&C included) have delved into higher-level quantum biology which now can explain, amongst other complex biological systems, navigation in some migratory birds. Quantum mechanics even extends to the populist level what with the Canadian PM being lauded in a recent presentation in which he explained (very sort of) the mechanics of quantum computing (*The Guardian*, 16 April 2016). Indeed, if we take this a step further, since language = thought requires unheard amounts of computing power, quantum computing may well provide some models for language as well. Obviously, I lack the technical expertise in these matters, but even when reading a popular text on quantum biology (McFadden & Al-Khalili 2014; see also my review of the book in Sperlich 2015), one can make quite a few interesting suggestions for language. For example, the famous linguistic bugbears of 'displacement' and long-distance binding in anaphora can be envisaged

as quantum states, that is, as remaining connected or intertwined over long distances. The many 'spooky' phenomena of quantum mechanics may play major roles in the neurophysiology of the brain and should be of interest to biolinguists as well.

In any case, I do hope that my plea for B&C serves to convince the jury of learned biolinguists of the merits of B&C's arguments, thus being able to further investigate language on the basis of mental Merge, generating thoughts (and the lexicon when such thoughts are externalized).

References

McFadden, Johnjoe & Jim Al-Khalili (2014). *Life on the Edge: The Coming of Age of Quantum Biology*. New York: Crown Publishers.

Berwick, Robert C. & Noam Chomsky. 2016. *Why Only Us: Language and Evolution*. Cambridge, MA: MIT Press.

Bolles, Edmund B. (2016). Chomsky on *Why only us*. Babel's Dawn blog (27 January 2016), http://www.babelsdawn.com/babels_dawn/2016/01/chomsky-on-why-only-us.html.

Evans, Vyvyan. (2016). *Why Only Us*: The language paradox. *New Scientist* 3062 (27 February 2016).

Fujita, Koji. (2009). A prospect for evolutionary adequacy: Merge and the evolution and development of human language. *Biolinguistics* 3: 128–153.

Hauser, Marc D. (2016). Challenges to the what, when, and why? *Biolinguistics* 10: 1–5.

Murphy, Elliot. (2016a). Review: 'Why Only Us' by Berwick and Chomsky. https://elliot-murphy.com/2016/01/16/review-why-only-us-by-berwick-and-chomsky.

Murphy, Elliot. (2016b). The human oscillome and its explanatory potential. *Biolinguistics* 10: 6–20.

Murphy, Elliot. (2016c). Phasal eliminativism, anti-lexicalism, and the status of the unarticulated. *Biolinguistics* 10: 21–50.

Sperlich, Wolfgang. (2015). Life on the edge. http://wolfgangsperlich.blogspot.co.nz/2015/08/life-0n-edge.html.

Wolfgang Sperlich
Northland Polytechnic
Department of English
155 Queen St.
Auckland
New Zealand
wsperlich@northtec.ac.nz