

---

**From:** Joscha Bach [REDACTED]  
**Sent:** Wednesday, March 1, 2017 7:29 PM  
**To:** Jeffrey Epstein  
**Cc:** Barnaby Marsh  
**Subject:** Re:

Thank you, Jeffrey! This is from Noam, right? I would be very interested in reading the responses of linguists and computational language modelers to this.

May I forward it to a friend at Google X?

Some notes:

> basic assumptions about human language that should I think be uncontroversial, extensively discussed elsewhere, then turning to a sample of challenges. A person's language is an internal system, part of human biology, based on configurations of the brain, developing in each person through interaction of specific biological endowment (the topic of UG – universal grammar in contemporary terminology), external environment, and general properties and principles of growth and development that are independent of language. =

As far as I understand, there is not yet an agreement among linguists on UG, i.e. how much is innate vs. do humans just converge on the simplest type 3 grammar that is consistent with the constraints they observe in their local environment. I think Noam argues that we have very specific circuitry for language, whereas the other camp would suggest that we are general learners, with specific rewards that bias us towards compositionality and systematicity. OTOH, this might also be read as a variant of Noam's "Strong Minimalist Thesis" (SMT).

The controversy will be eventually resolved by progress in building systems that learn natural language.

> The acquired system is an "internal language" (I-language), a computational system that yields an infinite array of hierarchically structured expressions that are interpreted at the conceptual-intentional CI interface as, in effect, a "language of thought" (LOT), and that can be externalized to one or another sensorymotor system, typically sound. Also relevant are some considerations about evolution of language.

>

> Little is known about the evolution of cognitive faculties, a matter discussed in an important article by Richard Lewontin, whose own view of the prospects was dim

Most folks in cognitive science would probably agree that most cortical activity is devoted to building a generative simulation of the outside world by a process of hierarchical learning. These simulations can be mapped on a conceptual manifold, something like an address space of our sensory motor representations of the world, which we can use to evoke and shape our mental simulations. Language is our interface to that conceptual space, and external language allows us to synchronize concepts even in the absence of matching sensory motor representations, i.e. we can build mental simulations of things that we never experienced, by interpolating between concepts that address mental simulations we know.

It seems that Noam's approach is unique in that he focuses entirely on language and concepts, while treating the understanding of the underlying cognitive faculties as hopeless, while many others would argue that understanding language without first understanding pre-linguistic mental representations might be impossible.

That said, Noam's characterization of I-language and LOT at the conceptual-intentional interface, with an externalization through generative mechanisms, is probably a useful basis, regardless of where individual researchers come from.

> [i] Anatomically modern humans (AMH) appear about 200 thousand years ago.

- >
- > [ii] The faculty of language FL appears to be a true species property: =hared among human groups (with limited individual differences) and in =ll essential respects, unique to humans. In particular, there is no =eaningful evidence for existence of language prior to AMH.
- >
- > [iii] Recent genomic studies indicate that some human groups (San =people) separated from other humans about 150kya. As far as we know, =hey share FL with other human groups.
- >
- > [iv] The San languages are all and only those that have the curious =property of phonetic clicks, and there may be some articulatory =adaptation to producing them (See Huijbregts, forthcoming).

Nguni languages have clicks, too, but they seem to have imported them =rom Khoisan.

> [v] The first (very limited) indication of some form of symbolic =behavior appears at about 75kya. Not long after that, we have rich =vidence of quite extraordinary creative achievements (Lascaux, etc.).

This is consistent with another observation: Modern humans had a =opulation bottleneck of 2000-3000 individuals ca 75000 years ago, which =coincides with the Tonga eruption. This does not necessarily mean that =he volcano killed off almost all hominids, but it increased the =volutionary pressure, and it is possible that our ancestors evolved a =utation that enabled them to outcompete and kill most of the hominid =ompetition (including Neanderthals). What if that mutation is something =hat roughly translates into "symbolic behavior"?

I currently think that much of our civilization might be the result of a =eries of quite specific mutations. Our ancestors went from 3000 =ndividuals to one million and remained there until they developed =eligions. Religion and other ideologies are based on a need for =onformance to internalized norms, i.e. an innate desire to serve as =art of a system that is larger than the individual's reputation based =roup. They were also based on a shared conceptual space.

Challenge 1 seems mostly to amount to: verify that 1. all human groups =ave language, and 2. there is no grammatical non-human language. One of =he interesting questions might be if dolphins have grammatical =language, another one concerns the limits of learning in non-human =rimates. The challenge is completely empirical.

Challenge 2 seems very exciting to me; I read it as: has language =ntrinsically linear order, or is that only imposed by the =sequentialization of articulation? Grammatical language has a tree =tructure, and the tree seems to be created probabilistically in the =istener, from a string of discrete symbols. Would natural language be =earnable without the constraints of sequentiality and discreteness?

Challenge 3: do we need externalization to learn and process language? I =ould suspect that an individual can play a language game against itself =ntil it converges on its own language, but it is not clear that humans =re among the class of individuals that can do that from scratch. Most =esearch suggests that there is a critical window in which we must pick =p our first language for perfect fluency, and there seems to be no =vidence of entirely individualistic acquisition/formation of a first =language. If that is true, is that a constraint of the way language =earning is implemented in the human brain, or a complexity constraint =ithin language itself?

It seems to be clear that learning a programming language changes the =ay we think, i.e. it provides evidence for a weak version of the Sapir =hof hypothesis. But that is not so much a constraint of =xternalization, but of the semantic structures addressed by the =language.

I imagine that pure work in a computer science lab can make some =nteresting progress on challenges 2 and 3.

Challenge 4: I don't understand enough about the context to see =he significance yet; I would think that once we have an SMT model of =language formation, we can learn additional operations that perform =operations on the generated

mental representation, based on arbitrary signals. This may require us to leave an approach that attempts to sandbox language from general cognition, but why would we want to constrain SMT based models by such a sandbox?

Challenge 5: Again, I don't understand enough of the context to understand why probabilistic interpretation cannot fill in the gaps. A probabilistic model will weight alternatives, and the binary Merge is the simplest, preferred case?

Challenge 6: The question of the structure of individual lexical items might require a perspective that integrates mental representations beyond language/SMT.

Challenge 7: Do semantic atoms refer to the external world ("referential doctrine")? – This seems to be quite clearly false; they refer to representations in the neocortex that are mutable and acquired through learning (structure or reinforcement) and inference.

Challenge 8: Noam seems to agree with my take on 7. How are semantic items acquired? – This challenge comes down to the general problems of learning and perception, i.e. pretty much everything in cognitive science outside of language! Challenge 8 seems to be designed by a rocket scientist who specializes in combustion chambers and leaves all other parts of getting the rocket to fly as an exercise to their grad student...

Challenge 9: Noam suggests that meaning must be derived from innate information, and wants to study universals between language to identify the innate bits. However, it is not clear if they do not stem from the properties of mathematics, i.e. there is a limited space of "useful simple axiomatic systems" that can be individually explored by learning systems. Kant attempted to describe this space, identified it as apriori and synthetic, and listed the basic structural categories that we would use to characterize the world. Sowa and a few others have made contributions to basic ontologies, and perhaps it is time to revisit Kant's project?

Challenge 10: Do music, planning, arithmetic stem from language, or do all result from a shared innovation of modern hominid brains? – Obviously, different answers in that space might be possible, for instance music could be a parasitic byproduct of rewards for discovering compositional representations that our brain needs to make us interested in learning grammar, while basic planning is independent, and complex planning needs language for structuring and operating on the conceptual space. This makes the question extremely general.

It also gives rise to the more general question of what exactly makes homo sapiens different from the other chimpanzees. I suspect that our brains are trained layer by layer, whereby each layer has a time of high plasticity during its primary training phase, then undergoes synaptic pruning, and has low plasticity later on. The duration of the training phases is regulated by genetic switches. Increasing the duration will extend infancy and childhood (i.e. increase the cost of upbringing), but give each layer more training data. Perhaps humans outperform other apes because they get a magnitude more training data before their brains lose infant plasticity, which results in dramatically better ability to generalize and abstract?

Challenge 11: Rare constructions can be understood by children, and thus there should be a mechanism to derive them from more simple rules, despite apparent evidence to the contrary, which should be explained away].

Challenge 12: Noam suggests that the complexity of most constructions in the face of "poverty of stimuli" means that I-languages are 1. very similar, 2. differences result from externalization, 3. should therefore stem from UG. He wants this shown, or an alternative.

An alternative explanation might be that the space of possible human grammars is small enough to allow rapid convergence, and in polyglots even allow for a complete mapping. That would not be a property of an evolutionary-engineering UG, but an apriori of the mathematics of human grammars.

Challenge 13: What small change in a brain could lead to the unique cognitive abilities of homo sapiens, including language? – There are a lot of different hypotheses of this, among them what I suggest in 10), and differential

attention/reward for learning compositional structures, or several successive modifications in the reward system. I think that Noam suspects that the culprit is a new connective pathway, perhaps somewhat similar to Julian Jayne's Bicameral Mind hypothesis?

These challenges are extremely inspiring food for thought!

Bests,

Joscha

> Am Mar 1, 2017 um 7:01 AM schrieb jeffrey E. <jeevacation@gmail.com>:  
>  
> <Challenges Language 2-17.docx>  
  
<?xml version=.0" encoding=TF-8"?>  
<!DOCTYPE plist PUBLIC "-//Apple//DTD PLIST 1.0//EN" "http://www.apple.com/DTDs/PropertyList-1.0.dtd">  
<plist version=.0">  
<dict>  
    <key>conversation-id</key>  
    <integer>52117</integer>  
    <key>date-last-viewed</key>  
    <integer>0</integer>  
    <key>date-received</key>  
    <integer>1488396562</integer>  
    <key>flags</key>  
    <integer>8590195717</integer>  
    <key>gmail-label-ids</key>  
    <array>  
        <integer>6</integer>  
        <integer>2</integer>  
    </array>  
    <key>remote-id</key>  
    <string>692309</string>  
</dict>  
</plist>