Towards eliminating arbitrary stipulations related to parameters: linguistic innateness and the variational model

Article

Accepted Version

Authors Accepted Version- Linguistic Approaches to Bilingualism

Towards Eliminating Arbitrary Stipulations Related to Parameters: Linguistic Innateness and the Variational Model

Jason Rothman[1,2] and Noam Chomsky[3,4]

University of Reading[1], UiT the Arctic University of Norway[2], Massachusetts Institute of Techonology[3], University of Arizona[4]

The Variational Model (VM) (Yang 2002, 2016) is inspired by evolutionary models of biological change; in a way, a grammar survives if indeed it is the fittest; "a grammar whose competitor is penalized more often will be learned faster" (Yang, 2018:XX). According to the VM, the developing learner simultaneously entertains a number of possible grammars associated with stochastic weighting. Underlying grammatical representations (intermediary and ultimately attained ones alike) are contingent on the child's tracking of probabilistic distributions in response to available input. In our view, Yang's approach is an exciting step forward for understanding the dynamics of language acquisition, not the least because it reconciles some seemingly contentious debates. Highlighting the indispensability of input for determining the ultimate system, it carves out a proper place for innate determination of the set of potential grammars and linguistic categories, along with a learning theory that may well be in part human-specific.

In the present paper, Yang (2018) describes and explains the mathematics behind the VM and its coverage of impressive data sets in child (monolingual) acquisition. The most novel part of the paper involves the application of the VM to the domain of adult second language acquisition, uncovering at the same time potential insights that could eventually be extended/modified to bilingualism more generally (e.g., heritage language bilingualism, L1 attrition, simultaneous child bilingualism). Given the complexities and range of topics touched upon, there is no shortage of points worthy of thoughtful commentary. We expect that several colleagues will focus on the relative potential for an (increasingly) successful application of Yang's model to bilingualism. And so, we will focus on something else that should be of ubiquitous concern to all who study language acquisition/attrition, be it monolingual, bilingual or multilingual in childhood or adulthood. We will focus on the following two interrelated topics: understanding (a) the Variational Model as a natural development of generative models of language acquisition and (b) how arbitrary stipulations of parameter values can be eliminated in terms of the learning theory he has developed.

The recognition that input plays an indispensable role for mental grammatical construction is shared by all cognitive based approaches to language acquisition, as is a (certain level of) innateness. Usage-based theories are fully innatist; the difference from a generative grammar view is the claim that what is innate is strictly domain-general. The investigation of generative grammar, when concerned with acquisition, has led to the conclusion that elements of domain-specificity have a crucial role in determining the hypothesis space (the set of potential grammars). In this context, it is useful to keep in mind that generative grammar is the theory of the possible states attained. As such, generative grammar itself is neutral about how the final state (a grammar/I-language, or probability distribution over them) is attained. It is beyond question that much of the resulting idiosyncratic properties of particular grammars, including the specificity/size of vocabulary and how we use language, depend on linguistic

exposure[1]. There is no principle of generative grammar about that, instead generative grammar is about the phenotype. It is, then, an empirical conclusion that the innate basis for language includes domain-specific principles (Universal Grammar, UG).

There has been substantial progress in determining what the innate principles might be, though much remains unknown. Nevertheless, they do appear to be unique to humans and crucial, even for word learning. Let's take, for example, the very simple principle that yields structure dependence, a core feature of language as distinct from any symbolic system known in the animal world, all of which seem to use the very simple properties of linear order (as other human cognitive processes do). The empirical question is to determine the principles that enter into language acquisition, and to determine which are language-specific. Yang's work is one of the rare contributions identifying what may be a more general learning principle, which relates to UG specifically showing that at least some valuation of parametric variation may not have to be stipulated for particular languages. In doing so, it leaves the UG framework (right or wrong) otherwise unchanged, eliminating previous stipulations proposed to account for descriptive facts/observations.

What better distinguishes generative approaches and usage-based/emergentist ones has never rested on whether or not the language acquisition device is context and/or frequency sensitive—it must be—but on whether some aspects of language are domain-specific, or none at all. Chomsky (1959) is often cited as the genesis of the generative linguistic enterprise, and so it is worth revisiting statements as they originally appeared:

> As far as acquisition of language is concerned, it seems clear that reinforcement, casual observation, and natural inquisitiveness (coupled with a strong tendency to imitate) are important factors, as is the remarkable capacity of the child to generalize, hypothesize, and "process information" in a variety of very special and apparently highly complex ways which we cannot yet describe or begin to understand, and which may be largely innate, or may develop through some sort of learning or through maturation of the nervous system. The manner in which such factors operate and interact in language acquisition is completely unknown. It is clear that what is necessary in such a case is research, not dogmatic and perfectly arbitrary claims, based on analogies to that small part of the experimental literature in which one happens to be interested (Chomsky 1959:43).

In 2018, almost 60 years on, several parts of the above quote could be updated although the essence remains. It is as undeniable today as it was in 1959 that "*[…] casual observation, and natural inquisitiveness (coupled with a strong tendency to imitate) are important factors".* We do know much more today about how factors mentioned above "*operate and interact in language acquisition*", although few would deny we have much more to uncover. Over the last six decades, various proposals have articulated opposing views on what underlies *"the remarkable capacity of the child to generalize, hypothesize, and "process information" in a variety of very special and apparently highly complex ways"* although we are now better equipped to describe and eventually understand them. The question of what subset of this

[1] Alignment with linguistic exposure is of course not a simplistic one-to-one mapping, even for word learning or aspects thereof (see Gleitman & Fisher, 2005; Lidz, Gleitman & Gleitman, 2003). For example, the complex meaning of even the simplest words, those used to refer, is acquired on very few exposures. Exposure seems to involve mostly Saussurean arbitrariness. In fact, there is no imaginable course of experience that could yield these meanings, revealed when one looks at them carefully.

capacity is "*largely innate or may develop through some sort of learning or through maturation of the nervous system*" is still an open one, although nearly sixty years of theorizing within generative grammar has narrowed the gap, reducing many specific assumptions about more general principles of UG and computational efficiency.

As Yang points out, the theory of innate parameters—the idea that UG provides a set of highly specific universal principles with constrained options to be set by the child—has been developed extensively since it was proposed 40 years ago. There has been substantial work on the nature and organization of parameters, on ways to reduce the search space in acquisition, and in recent years, on ways that parameters might be reduced to principles of efficient computation (e.g., Berwick & Chomsky, 2016; Epstein, Obata, & Seely, 2017). Parameters in the classic sense were (and remain) a fundamental construct. At the conceptual level, the theory offered the first suggestion about a feasible acquisition system, with bounded choices. It led to inquiry into languages of scale and depth that goes beyond anything in the history of the subject, along with very fruitful work on the nature and structure of linguistic variation. Its legacy includes fine-grained descriptions of grammatical domains, a program designed to describe and explain linguistic variation and observations of the co-occurrence of grammatical phenomena that otherwise would not be expected (see e.g., Fábregas, Mateu, & Putnam, 2015; Karimi & Piattelli-Palmarini, 2017 *inter alia*). It led to new work in diachronic linguistics, previously unformulable, about setting of parameters and how it changed (Lightfoot & Westergaard, 2007). It injected innovation into psycholinguistics, both processing and acquisition (see Hyams et al., 2015; Sanz, Laka, & Tanenhaus, 2013). The granularity of the theory lent itself well to making order-of-acquisition predictions, explaining patterns of production, providing a toolkit to describe developing language in formal terms, as well as offering a seemingly good-enough answer to the logical problem of acquisition. Related to the Chomsky (1959) quote, it was a proposal on what "may be largely innate".

Yang's work doesn't bear on the very existence of parameters but rather on how they are valued. There are three questions to face about parameters: (1) what they are: the options available for language (2) how they are set in particular languages and (3) where they come from[2]. Yang's work deals with (2), essentially assuming (1). Let's consider the work on double objects. He assumes that there are two options: double object/(dative) to-phrase, which constitutes the parameter. Innumerable other ways of organizing the (limited) data available to the child are ignored. The Tolerance Principle very insightfully shows how the setting can be determined (caused possession, etc.). This eliminates the unwelcome need to stipulate values of parameters. But it doesn't eliminate the parameter per se, which determines the options, excluding innumerable others. Here we see the greatest contribution of Yang's model: its overcoming the stipulation of values for innate linguistic parameters. In doing so, Yang's work focuses on gradual yet highly constrained development. His algorithmic approach provides

[2] Given that it is hard to imagine parameters evolved, serious work regarding (3) is in progress. Take the head parameter. If nature selected the simplest way to yield discrete infinity (Merge), then the structures generated internally, and semantically/pragmatically interpreted, have no linear order. But the articulatory system, which is independent of language, requires linear order. So there is a mismatch between two unrelated systems (it's a bit different with sign, which has more dimensions available). Externalization must assign linear order, and there are two ways, head-first/-last. But there is no parameter. Samuel Epstein and his group have developed such ideas further, discussing cases of variation where, they argue, it is a matter of a choice that has to be made in rule ordering where UG leaves it open. Such approaches actually do get rid of parameters – but what lies ahead is an immense task, posed clearly by the fundamental work of Mark Baker, Giusseppe Longobardi, Richard Kayne, and many others.

.

specific answers to the question when input data yield rule formation rather than listing, with quite remarkable results.

Yang's work shows very effectively how unlearned/unlearnable properties of language interact with specific linguistic experience in child language acquisition, a very significant contribution. With Yang, we hope that the VM and specifically the Tolerance Principle will enable further progress in understanding these interactions in the many domains that are yet to be explored. Some twenty years on, it is still true today that:

> The field is changing rapidly under the impact of new empirical materials and theoretical ideas. What looks reasonable today is likely to take a different form tomorrow […] Though the general framework remains, the modifications at each point are substantial […] The end result is a picture of language that differs considerably from its immediate precursors. Whether these steps are on the right track or not, of course, only time will tell (Chomsky, 1995:9).

References

Berwick, R. C., & Chomsky, N. (2016). *Why only us: Language and evolution*. Boston. MIT press.

Chomsky, N. (1959). Verbal behavior (a review of Skinner's book). *Language*, *35*, 26-58.

Chomsky, N. (1995). *The Minimalist Program.* Boston. MIT Press.

Epstein, S. D., Obata, M., & Seely, T. D. (2017). Is linguistic variation entirely linguistic? *Linguistic Analysis*, *41*(3), 481-516.

Fábregas, A., Mateu, J., & Putnam, M. (eds.). (2015). *Contemporary Linguistic Parameters*. New York, NY. Bloomsbury Publishing.

Gleitman, L., & Fisher, C. (2005). Universal aspects of word learning. In McGilvray, J. (ed.) *The Cambridge Companion to Chomsky*. Cambridge, Cambridge University Press.

Hyams, N., Mateu, V., Orfitelli, R., Putnam, M., Rothman, J., & Sánchez, L. (2015). Parameters in language acquisition and language contact. In Fábregas, A., Mateu, J., & Putnam, M. (eds.). (2015). *Contemporary Linguistic Parameters*. New York, NY. Bloomsbury Publishing.

Karimi, S. and M. Piattelli-Palmarini (eds.) (2017). Parameters: What are they? Where are they? Special issue, *Linguistic Analysis*, Vol 41 Issue 3 & 4.

Lidz, J., Gleitman, H., & Gleitman, L. (2003). Understanding how input matters: verb learning and the footprint of universal grammar. *Cognition, 87*(3), 151-178.

Lightfoot, D., & Westergaard, M. (2007). Language Acquisition and Language Change: Inter-relationships. *Language and Linguistics Compass*, *1*(5), 396-415.

Sanz, M., Laka, I., & Tanenhaus, M. K. (eds.). (2013). *Language Down the Garden Path: The Cognitive and Biological Basis for Linguistic Structures*. OUP Oxford.

Yang, C. (2002). *Knowledge and learning in natural language*. Oxford. Oxford University Press on Demand.

Yang, C. (2016). *The price of linguistic productivity: How children learn to break the rules of language*. Boston. MIT Press.

Yang, C. (2018). A Formalist Perspective on Language Acquisition. *Linguistic Approaches to Bilingualism, 8(6)*