Making sense of syntax – Innate or acquired? Contrasting Universal Grammar with other approaches to language acquisition

Christian Kliesch
School of Psychology
University of Glasgow

Manuscript submitted for publication. Please do not quote.

Proponents of a Universal Grammar approach argue that humans are born with a dedicated language system that shapes and restricts the number of grammars found in human languages (Chomsky, 2005). It is essentially innate and has a genetic manifestation. Such an innate system is necessary because human grammars are too complex to be passed on through social interactions and probabilistic learning alone. However, this view is currently contested by a combination of emgergentist approaches and a number of studies suggest that many of the core assumptions of Universal Grammar are either unnecessary or do not hold. Furthermore, this review will explore theoretical criticism of the Universal Grammar research programme.

Keywords: Universal Grammar, Language Acquisition, Language Evolution, Poverty of Stimulus, Nativism

Learning to understand a language is a task of seemingly incredible complexity, which appears to be almost infinitely complex from an infant's perspective. But the fact that infants are indeed capable of, and in fact do, learn a language almost effortlessly suggests that this is not the case. Currently, there are a number of different approaches that try to explain the processes used by infants to decipher language by acquiring new words and learning the grammar of their language. The predominant approach of the last 50 years has been that children have an innate sense for grammar and syntax, shared by all humans—a Universal Grammar (Chomsky, 2005; V. Evans & Green, 2006). The universal grammar approach is part of Chomsky's generative view of language acquisition (V. Evans & Green, 2006). The initial paradigm by Chomsky was a response to Behaviourism, which could not fully explain acquisition in children. Human grammars are too complex to be learnt by learning processes alone (Chomsky, 2005). Furthermore, a wide variety of languages investigated, suggested that their syntaxes share common features and traits, even though they were not directly related. The Behaviourist view that linguistic acquisition was free and independent, whereas grammar could take any possible form, seemed refuted (Chomsky, 2005), and language acquisition seemed best explained by an innate Universal Grammar. However, this view has recently been challenged by a number of other approaches, which combined, form a strong alternative to an innate syntax module, as predicted by Universal Grammar theorists. They attack the main arguments of Universal Grammar, mainly that (1) there is an innate language area in the brain responsible for syntax processing (2) which restricts the expression of possible grammars (this will lead to syntactical language universals) and (3) is genetically codified. Furthermore, (4) the stimulus presented in language acquisition is too poor to allow acquisition through interactional and social constrains (Hauser et al., 2002).

Universal Grammar is a biolinguistic approach to language acquisition and usage (Chomsky, 2005). The language model proposed by nativists is domain specific; the cognitive processes underlying language processing differ qualitatively from other learning processes used by the child to understand the environment(Chater & Manning, 2006). According to this view the brain is consists of different 'organs', or modules, of which one is responsible for syntactical features. However, current nativist models vary with regard to what is actually predisposed in language. Whilst early approaches to universal grammar predicted that all languages would share specific syntactical features, later revisions of Chomsky's theory argue that Universal Grammar would serve as a preselector for all the available grammars, and depending on the socio-linugal context, the appropriate syntaxes would be acquired by the infant (Yang, 2004; Chomsky, 2005). Hence, Universal Grammar restricts possible human syntaxes, but still allows room for variation. Thus, it is argued, that child explores possible options and gradually tunes it to the adults (Crain et al., 2006). In support of this argument Crain et al. (2006) point to the linguistic continuity hypothesis: There are certain grammatical constructs common to all languages. Whilst some are correct in some languages, children would initially use these constructs even in languages that do not

Contact e-mail address: ck [at] antipattern.net

have these features. Through social learning and imitation they would discard such false constructs over time. At the same time, there are some possible constructs which are logically sound (i.e. can be deduced from experience) but are not used by children. Crain et al. (2006) have found evidence for such patterns in a comparison of children's language errors in German, English and Italian why questions, as well as disjunctions in child Japanese. Thus it was found that one third of children of native English speakers in the sample, inserted an extra medial-wh in long-distance questions ("What do you think what pigs eat?", p 32) which is very similar to the structure actually used in German. There is similarity of false why-question inversion (e.g. "Why that's not your sandwich?", p 35) and the question structure in other languages, for example Italian. Furthermore they found that Japanese children interpret Syntactic Subsets in a way that English speakers do, but that is not used in adult-Japanese (Crain et al., 2006). Thus the current minimalist paradigm is an advance over the previous approaches to universal grammar. Instead of expressing clear rules for grammatical the minimal paradigm suggests that rules are of an abstract nature. Earlier Universal Grammar were much more specific in their syntactical predictions (N. Evans & Levinson, 2009).

In fact, N. Evans & Levinson (2009) argue that any universals are over-generalisations. They note that Universal Grammar theorists overlook the enormous linguistic diversity. Currently there are 5000–8000 different human languages, of which only 10 % described in grammar and dictionary and it is estimated that there existed 500,000 human languages over time. However, universal grammar research draws mostly on well-known languages, and hence N. Evans & Levinson (2009) argue that no inferences can be drawn from so little data. Furthermore there is only very little data from isolated languages, of which some lack even simple features of morphology, or do not use adverbs and adjectives. Others have features like ideophones, positionals and coverbs (N. Evans & Levinson, 2009), unknown to western languages. Therefore, N. Evans & Levinson (2009) argue that humans are *"the only known species whose communication system varies fundamentally in both form and content."* [p. 431]

According to the nativist view of Universal Grammar, these achievements would be impossible without an innate and universal grammar: The language stimuli that children are exposed to are simply too noisy and incomplete to allow for reliable understanding of language through induction. Thus, much of the support for an innate universal grammar is drawn from the observation that children are exposed to impoverished language stimuli. Whilst Universal Grammar allows for probabilistic and social learning to acquire language-specific rules, it states that this would not be possible without an innate knowledge of the syntactic and phonological structure (Yang, 2004; Chomsky, 2005). Furthermore, defenders of the poverty of stimuli argument state that parents do not give corrective feedback on grammatically false utterances, particularly in early childhood (Brown & Hanlon, 1970). However, recent studies by Strapp et al. (2008); Saxton (2000); Saxton et al. (2005) present evidence that there is in fact considerable, albeit indirect, negative feedback if a child uses the wrong grammar. Thus Strapp et al. (2008) found that children learnt irregular forms of artificial verbs and noun plurals through negative evidence, and interestingly there was an interaction of word form and age: In 3 year old infants, negative evidence proved to be more effective for noun plurals, whilst 5 year olds used it for verbs more than nouns. This gives compelling evidence that negative evidence exists, and stimuli are not as impoverished. Although this does not refute the poverty of stimuli argument altogether, it is weakened considerably since it no longer falsifies theories arguing against an innate syntactical structure.

The poverty of stimulus argument is also being questioned by Chater & Christiansen (2009). They distinguish two kinds of inductive learning: (1) N-induction, the ability to understand a seemingly arbitrary external world. (2) C-induction, the ability to co-ordinate with others. Whilst the former does indeed pose an impoverished stimulus to the child, the latter only requires to do as others do. The necessary skills for C-induction do not need to reside in a 'central grammar organ' but are to be found in innate social skills. They further elaborate that N-induction is fairly stable, whilst C-induction shows high variability and depends on the cultural context.

Social approaches may be based on innate principles, too. But these are much more general than a universal grammar (Seidenberg, 1997). Furthermore, a concept of language acquisition though social interactions can be found in other species as well (Goldstein & Schwade, 2009; Kuhl, 2004). Thus, cowbirds do not grow up with members of their own species, it was assumed to have an innate predisposition for species-specific song. However, this assumption no longer holds: They acquire song through feedback from female partners. If cowbirds are not exposed to other birds, the will not acquire full song (Goldstein & Schwade, 2009). Similar reports come from deaf and hearing impaired infants, who develop different babbling to children who do not have hearing impairments (Goldstein & Schwade, 2009).

A central part in the debate for and against a universal grammar is the whether or not a child is capable of inducing a grammatical structure from the input. However, proponents of a social view of language acquisition suggest that social cues give sufficient infant-directed speech has certain properties that facilitate language learning (Uther et al., 2007; Goldstein & Schwade, 2009; Saxton, 2000; Saxton et al., 2005; Kuhl, 2004). A comparison of infant directed speech and foreigner directed speech by Uther et al. (2007) has shown that the both share features like vowel-hyperarticulation important in making sense of language whilst differing in pitch and rated affect, which are thought to satisfy children's socio-emotive needs. Thus they accommodate the shared linguistic need of both groups, whilst differing in social-emotive needs. Unfortunately, Uther et al's study looks at word acquisition, a comparat-

ive approach to the similarities of grammatical features has not yet been done. Additional evidence that child-directed speech has characteristics that support a child's language acquisition is presented by Brodsky et al. (2007) who emphasise the importance of partial repetition of phrases by parents. Parents use a high number of phrase repetitions when interaction with a child. Repetitions alone would not give sufficient variation of input. However, parents offer the child a series of *variation sets*, parents repeat utterances with small changes, leading to changes in the syntactical structure but not the content. For example (taken from Brodsky et al. (2007), p. 2):

```
You got to push them to school.
Push them.
Push them to school.
Take them to school.
You got to take them to school.
```

Such variations would be ideal for the child's word and syntax acquisition process. And indeed, Küntay & Slobin (1996) found that they make up up tho 20 % of child-directed speech in Turkish. In an analysis of children's speech, it was found that there is a high correlation with parents variation sets and subsequent language and syntax use in children (Brodsky et al., 2007). In a quantitative analysis of motherease Brodsky et al. (2007) were able to construct an algorithm capable of acquiring a relatively high-coverage of generative grammar. Such probabilistic models achieve 54 % precision on Mandarin and 63 % on English, with an approximate recall of 30 % for both—using child-directed speech as an input (Waterfall & Edelman, 2009). Based on these findings, Waterfall & Edelman (2009) suggest that some of the grammatical features are simply more common because they are more readily acquired by probabilistic learning processes. Taken together, this evidence suggests that social interactions play a major role in language acquisition, a view for which early approaches to Universal grammar only left little room (Chomsky, 2005). This is different to recent revisions of Universal Grammar theory, which consider social input as part of the 'tuning' process used by children to choose the right grammar from the ones that are available. Hence, whilst these social-interactional findings reduce the necessity for Universal Grammar, they do not refute it either.

Universal Grammar is assumed to be expressed through genes. A distinct linguistic system would have evolved through natural selection, for example favouring a mutation that lead to the rewiring of neural networks to accommodate basic syntactical structures (Chomsky, 2005). Criticism comes from Chater & Christiansen (2009) who argue that—given the short timeframe of language evolution—the development of a genetically codified language module is improbable. Unlike dedicated systems like the visual-perceptive system, a genetically encoded universal grammar would require much more complex mutations. Even though nativist approaches highlight the similarities and shared features of different languages across all spectra, languages are too variable to allow for the evolution of a universal grammar, in the words of (Christiansen et al., 2009, p 221) the observed variability would pose "a moving target" for natural selection, attempts to replicate the development of a domain-specific module in computer simulation failed (Christiansen et al., 2009). Instead of a dedicated system that developed through natural selection, it is more likely that the reverse holds: Language has adapted to predating constraints posed by the brain, neural networks and already existing learning structures. The process of grammar development would be one of cultural evolution, within biological constraints (Christiansen et al., 2009; Christiansen & Chater, 2009; N. Evans & Levinson, 2009). There may be the possibility that there are initial heuristics used to parse language input and are adapted over time, but these are no constraints in the way that universal grammars are.

Support for a general learning processes comes from a study by Saffran et al. (2007) who replicated an earlier study by Marcus et al. (2004). Using triad-sequences of stimuli, the original study found that 7-month old infants were able to generalise linguistic stimuli, they failed to do so for shapes and non-linguistic sounds, like geometric shapes. However, Saffran et al. (2007) were able to show that infants are indeed capable of doing so, but may categorise stimuli differently, and thus not perceiving the stimuli presented by Marcus et al. (2004) as sequences. However, when presented with stimuli they know well and show interest in (for example dogs and cats), they are indeed capable of distinguishing between these sequences. Nevertheless, these triads represent very basic rules only, and further research is necessary to generalise these findings to more complex syntaxes

From the evidence considered in this review, it seems unlikely that the theory of hard-coded Universal Grammar holds. The criticism voiced in N. Evans & Levinson (2009) is even more fundamental: They argue that specific approaches to universal grammar are already falsified by the variability observed or have become so general that they escape possible falsifications. But according to Popper (2002) a scientific hypothesis must be potentially falsifiable in order to be considered scientific. But as N. Evans & Levinson (2009) state, "Chomsky's notion of Universal Grammar (UG) has been mistaken, not for what it is – namely, the programmatic label for whatever it turns out to be that all children bring to learning a language – but for a set of substantial research findings about what all languages have in common." [p 430] Its current character is therefore not of explicative value, but descriptive only.

Children show many instances of statistical and social learning in other instances of language acquisition. This shows that they are indeed capable of using these devices appropriately and drawing correct inferences. Thus the necessity for an innate universal grammar for explaining how children acquire language is low. This does not rule out that there are indeed some innate features of language. But it is more likely that these originate from cognitive, socio-interactional and probabilistic constrains (Seidenberg, 1997; N. Evans & Levinson, 2009). Tomasello (2009)

argues that these could potentially fulfil the criteria for a very broad concept, but that historically the term Universal Grammar requires biological manifestations of syntax of one sort or another: "It is not the idea of universals of language that is dead, but rather, it is the idea that there is is a biological adaptation with specific linguistic content that is dead." (Tomasello, 2009, p. 471)

Based on Imre Lakatos' epistemological approach (Lakatos & Feyerabend, 1999), one could argue that in its beginnings Universal Grammar approach was a *progressive* research programme that sparked new debates in the field of syntax acquisition and exposed weaknesses in previous research (Chomsky, 2005). However, currently the Universal Grammar programme offers little new input to the debate of how grammar is acquired, and one could argue that it is a *degenerating* research programme. Its explanative value regarding language acquisition is limited and well contested by other approaches. At the same time one of its key concepts—shared syntactical features—shows greater variability than predicted. Attempts to resolve these problems have watered down the potential falsifiability and predictability of its hypotheses (N. Evans & Levinson, 2009). But at the same time, combinations of evolutionary, socio-interactional and probabilistic approaches have not been able to explain the full picture either. Whilst combinations of probabilistic and social learning, as well as evolutionary accounts are promising, there is not yet a consistent framework that explains language acquisition as a whole, as well as the role that each of these models plays within the whole system.

References

Brodsky, P., Waterfall, H. R., & Edelman, S. (2007). Characterizing motherese: On the computational structure of child-directed language. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th cognitive science society conference* (pp. 833–838). Amsterdam.
3

Brown, R., & Hanlon, C. (1970). Cognition and the development of language. In J. R. Hayes (Ed.), *Proceedings from Fourth Annual Symposium on Developmental Linguistics, Carnegie-Mellon University, Pittsburgh Pa., April 11–12, 1968* (p. 11-53). New York: Wiley.
2

Chater, N., & Christiansen, M. (2009). Language acquisition meets language evolution. *Cognitive Science*, *1*, 1–27.
2, 3

Chater, N., & Manning, C. (2006). Probabilistic models of language processing and acquisition. *Trends in Cognitive Sciences*, *10*(7), 335–344.
1

Chomsky, N. (2005). Three factors in language design. *Linguistic Inquiry*, *36*(1), 1–22.
1, 2, 3, 4

Christiansen, M., & Chater, N. (2009). The myth of language universals and the myth of universal grammar. *Behavioral and Brain Sciences*, *32*(05), 452–453.
3

Christiansen, M., Chater, N., & Reali, F. (2009). The biological and cultural foundations of language. *Communicative & Integrative Biology*, *2*(3), 221–222.
3

Crain, S., Goro, T., & Thornton, R. (2006). Language acquisition is language change. *Journal of Psycholinguistic Research*, *35*(1), 31–49.
1, 2

Evans, N., & Levinson, S. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, *32*(05), 429–448.
2, 3, 4

Evans, V., & Green, M. (2006). *Cognitive linguistics: An introduction.* Edinburgh, UK: Edinburgh University Press.
1

Goldstein, M., & Schwade, J. (2009). From birds to words: Perception of structure in social interactions guides vocal development and language learning. In M. Blumberg, J. Freeman, & S. Robinson (Eds.), *Oxford handbook of Developmental Behavioral Neuroscience* (pp. 708–729). USA: Oxford University Press.
2

Hauser, M., Chomsky, N., & Fitch, W. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, *298*(5598), 1569–1579.
1

Kuhl, P. (2004). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience*, *5*(11), 831–843.
2

Küntay, A., & Slobin, D. I. (1996). Listening to a Turkish mother: Some puzzles for acquisition. In A. K. D. I. Slobin J. Gerhardt. & J. Guo (Eds.), *Social interaction, social context, and language: Essays in honor of Susan Ervin-Tripp* (pp. 265–286). Mahwah, NJ: Lawrence Erlbaum Associates.
3

Lakatos, I., & Feyerabend, P. K. (1999). *For and Against Method: Including Lakatos's Lectures on Scientific Method and the Lakatos-Feyerabend Correspondence.* Chicago, US: University Of Chicago Press.
4

Marcus, G., Johnson, S., Fernandes, K., & Slemmer, J. (2004). Rules, statistics and domain-specificity: Evidence from prelinguistic infants. In *29th Annual Meeting of the Boston University Conference on Language Development.* Boston, MA.
3

Popper, K. R. (2002). *Conjectures and refutations.* London: Routledge.
3

Saffran, J., Pollak, S., Seibel, R., & Shkolnik, A. (2007). Dog is a dog is a dog: Infant rule learning is not specific to language. *Cognition*, *105*(3), 669–680.
3

Saxton, M. (2000). Negative evidence and negative feedback: Immediate effects on the grammaticality of child speech. *First Language*, *20*(60), 221–252.
2

Saxton, M., Backley, P., & Gallaway, C. (2005). Negative input for grammatical errors: Effects after a lag of 12 weeks. *Journal of Child Language*, *32*(03), 643–672.
2

Seidenberg, M. (1997). Language acquisition and use: Learning and applying probabilistic constraints. *Science*, *275*, 1599–1603.
2, 3

Strapp, C., Bleakney, D., Helmick, A., & Tonkovich, H. (2008). Developmental differences in the effects of negative and positive evidence. *First Language*, *28*(1), 35–53.
2

Tomasello, M. (2009). Universal grammar is dead. *Behavioral and Brain Sciences*, *32*(05), 470–471.
3, 4

Uther, M., Knoll, M., & Burnham, D. (2007). Do you speak E-NG-LI-SH? A comparison of foreigner-and infant-directed speech. *Speech Communication*, *49*(1), 2–7.
2

Waterfall, H., & Edelman, S. (2009). The neglected universals: Learnability constraints and discourse cues. *Behavioral and Brain Sciences*, *32*(05), 471–472.
3

Yang, C. (2004). Universal grammar, statistics or both? *Trends in Cognitive Sciences*, *8*(10), 451–456.
1, 2