

Universal Grammar: Wittgenstein versus Chomsky

Danièle Moyal-Sharrock

In memoriam Laurence Goldstein (1947-2014)

A few years ago, I sent Laurence Goldstein a draft entitled 'Coming to Language: Wittgenstein's Theory of Language Acquisition'. Laurence being a Wittgenstein-inspired philosopher, I was astonished when his comments revealed a leaning toward Chomsky's Universal Grammar:

There is one problem that you mention but don't much discuss, about which I still feel some unease, and that's the 'poverty of stimulus' argument. You deny that there is any such poverty – you talk about the child's 'multifarious practice and repeated exposure', but child developmentalists say that infants are typically exposed to very little language and close to zero correction of grammar by parents. I am also disinclined to ignore Derek Bickerton's evidence for the 'language bioprogram hypothesis'. Bickerton gathered a large amount of data on pidgins and creoles. A pidgin has rudimentary grammar; a creole is grammatically complex, but the transition from one to the other is made within the space of one generation, suggesting that grammar is biologically hard-wired.

However, a week later, Laurence wrote me the following:

For the last week, I've been hanging around with my first grandchild, now six months old, and so have had the opportunity to assess the poverty of stimulus hypotheses. Of course, that environment, replete with articulate adults bent on amusing the child was unrepresentative. But what struck me, and this would be true too of the linguistically less rich environments, was the variety of 'language-games' to which the child is exposed. Almost all the words it hears are interwoven with action – objects are pointed to, animal sounds are made in the context of stories about country life, the child is lifted and lowered to the accompaniment of 'up we go.....down we go' etc.

In saying this, Laurence had replaced the poverty of *grammatically complex* instruction and correction with the richness of exposure to a huge variety of *language-games* where words, behaviour, context and repetition interact with each other to inculcate in a child her native language.

In this paper, I begin by unravelling some strands of the nativist argument, offering replies as I go along. I then give an outline of Wittgenstein's view of language acquisition to see if it doesn't render otiose problems posed by nativists like Chomsky, not least by means of Wittgenstein's own brand of grammar which, unlike Chomsky's, does not reside in the brain, but in our practices.

1. Chomsky's Universal Grammar: the nativist argument

... we humans have explicit and highly articulate linguistic knowledge that simply has no basis in linguistic experience.
Chomsky (1983)

The motivations for the claim that language is innate are, for many, quite straightforward. The innateness of language is seen as the only way to solve the so-called 'logical problem of language acquisition' (LPLA): the mismatch between linguistic input and linguistic output. How is it that children come to know and use – at an incredible speed – linguistic principles they have never been taught (and indeed, that exceed the knowledge of a PhD in linguistics), and how is it they can produce an unlimited number of sentences from the limited data they are exposed to? This is also known as 'poverty of the stimulus' or the underdetermination of the output. The nativist solution to this problem is that linguistic principles do not have to be input or learned at all; we are born with them – they come in the form of an innate Universal Grammar. For Chomsky, then, knowledge of language is based on a core set of principles embodied in all languages¹ and innately stored somewhere in the mind/brain of every human being. Let's flesh out the nativist argument.

The syntax or structure of any language is so abstruse that it seems impossible that children should learn it – particularly as quickly as they do. As Green and Vervaecke write:

Constituent hierarchical structure, an almost definitional feature of language, is just not something, by and large, that we come up against in the everyday world; and even when we do, it is darn hard, even for the best and brightest among us, to figure it out. Witness, for instance, the struggles of linguists themselves to adequately characterize language. ... linguists have been unable to discover exactly what the rules are, even after dozens (one might argue hundreds or even thousands) of years of research. By contrast, virtually every child does it within a few years (with far less in the way of specialized cognitive machinery, and control over the quality of the incoming data, it is worth pointing out, than a Ph.D. in linguistics). (1997)

The difficulty is compounded by the fact that the child's environment is, allegedly, of hardly any help. As Anderson and Lightfoot note: 'The child masters a rich system of knowledge without significant instruction and despite an impoverished stimulus; the process involves only a narrow range of 'errors' and takes place rapidly, even explosively between two and three years of age. The main question is how children acquire so much more than they experience' (2000, 13-14).

¹ In fact, nativists recognize that not all principles occur in every language, but claim that this does not prevent that principle from being universal as long as the principle is *not broken*. Indeed a principle can be claimed universal on the basis of its occurrence in a single language: 'In what sense can a universal that does not occur in every language still be universal? Japanese does not *break* any of the requirements of syntactic movement; it does not need locality for question movement because question movement itself does not occur. Its absence from some aspect of a given language does not prove it is not universal. Provided that the universal is found in some human language, it does not have to be present in all languages'; '... it is not necessary for a universal principle to occur in dozens of languages. ... it can be claimed to be universal on evidence from one language alone; I have not hesitated to propose a general principle of linguistic structure on the basis of observations of a single language' (Chomsky 1980b, 48)' (Cook & Newson 2007, 21; 23).

The *poverty of the stimulus* argument strikes at empirical or social theories of language acquisition by claiming that the utterances encountered by the child in experience are too limited² for it to be possible to learn the language by generalizing from them, and so we are forced to suppose that the brain contains innate means of creating an unlimited number of grammatical sentences from a limited vocabulary. Hence, Chomsky's stipulation that the child is born with a 'language acquisition device' (LAD) which, when the child starts being exposed to language, recognises which language it is and sets the correct parameters for that particular language. Thanks to the LAD, the child knows *intuitively* that there are some words that behave like verbs, and others like nouns, and that there is a limited set of possibilities as to their ordering in a sentence. The LAD can enable this because it is equipped with a Universal Grammar (UG) which consists of invariant principles³, as well as parameters⁴ whose settings vary between languages, and recursive rules to enable productivity or creativity. Thus equipped, the child is able to apply her built-in unconscious knowledge of how language works to the limited number of sentences she hears, and at an otherwise (allegedly) unexplainable speed⁵: 'Learning a particular language thus becomes the comparatively simple matter of elaborating upon this antecedently possessed knowledge, and hence appears a much more tractable task for young children to attempt' (Cowie 2008).

Minimal exposure to 'language evidence' is necessary to trigger the various parameters of Universal Grammar⁶ (Cook & Newson 2007, 186). As for vocabulary, writes Chomsky:

² Chomsky is no longer concerned by the degeneracy of the data, but only its poverty or meagreness. The poverty of stimulus argument now focuses on the poverty of language addressed to children (the fact that it does not contain the right kind of syntactic evidence) rather than on the degeneracy of the data (the fact that it is not always completely well-formed). This change is due to research on speech addressed to children which showed that it was highly regular, and so the data are arguably not as degenerate as was earlier thought. Newport *et al* (1977) found that only 1 out of 1500 utterances addressed to children was ungrammatical (Cook & Newson, 2007, 192-3).

³ UG is 'the sum total of all the immutable principles that heredity builds into the language organ. These principles cover grammar, speech sounds, and meaning' (Chomsky 1983); they are the finite, invariant, genetically-innate set of principles common to all languages 'by which the child can infer, on the basis of the limited data available in the environment, the full grammatical capacity which we think of as a mature speaker's knowledge of a language' (Anderson & Lightfoot 2000, 6). UG is part of the LAD, an innate biologically-endowed language faculty. The LAD is also known as the 'initial state' of the language faculty – the state we are born with; we have learned English (i.e. the language faculty reaches its 'mature state') when, by being exposed to it, we have learned the lexicon and set the parameters for English.

⁴ This is the Principles and Parameters (P&P) Theory, according to which 'UG provides a fixed system of principles and a finite array of finitely valued parameters' (1995, 170). Parameters are language-specific, binary parameters that can be set in various ways. An example of a parameter is 'the head parameter', whereby a particular language consistently has the heads on the same side of the complements in all its phrases, whether head-first or head-last. So, for instance, English is head-first: *in the house*: preposition head first before the complement; *killed the man*: verb head first before the complement. Japanese is head-last. 'It may be that the values of parameters are set to defaults at birth, but that these can be changed across a small range of values by certain linguistic experiences' (Green and Vervaecke 1997).

⁵ Bishop (2014) objects: 'The problem is then to explain how children get from this abstract knowledge to the specific language they are learning. The field became encumbered by creative but highly implausible theories, most notably the parameter-setting account [see note 4 above], which conceptualised language acquisition as a process of "setting a switch" for a number of innately-determined parameters'. I would, however, begin by objecting to the 'abstract knowledge'.

⁶ Anderson & Lightfoot: 'the trigger experience, which varies from person to person ... consists of an unorganized and fairly haphazard set of utterances, of the kind that any child hears' (2000, 14).

You just have to learn your language's vocabulary. The universal grammar doesn't tell you that "tree" means "tree" in English. But once you've learned the vocabulary items and fixed the grammatical parameters for English, the whole system is in place. And the general principles genetically programmed into the language organ just churn away to yield all the particular facts about English grammar⁷. (1983)

It is, then, through the *interaction* between our genetically-inherited principles and the linguistic environment to which we happen to be exposed that a specific language emerges:

... English-speaking children learn from their environment that the verb *is* may be pronounced [iz] or [z], and native principles prevent the reduced form from occurring in the wrong places. (Anderson & Lightfoot 2000, 6).

Let's see how this prevention works in practice. Anderson and Lightfoot:

The verb *is* may be used in its full form or its reduced form: English speakers can say either *Kim is happy* or *Kim's happy*. However, certain instances of *is* never reduce: for example, the [is] underlined items in *Kim is happier than Tim is* or *I wonder where the concert is on Wednesday*. Most speakers are not aware of this, but we all know subconsciously not to use the reduced form in such cases. How did we come to know this? As children, we were not instructed to avoid the reduced form in certain places. Yet, all children typically attain the ability to use the forms in the adult fashion, and this ability is quite independent of intelligence level or educational background. Children attain it early in their linguistic development. More significantly, children do not try out the non-occurring forms as if testing a hypothesis, in the way that they "experiment" by using forms like *goed* and *taked*. The ability emerges perfectly and as if by magic. (2000, 3)

On the nativist view, then, the child is faced with a chaotic linguistic environment and scans it – in this case, she is looking for *clitics*: unstressed words that cannot stand on their own (e.g., The contraction of *is*, in 'What's going on?' or the possessive marker 's in 'The man's book'). Since clitics and their behavior are predefined at the genetic level, the child is able to arrive at a 'plausible analysis' on exposure to a few simple expressions: she concludes that no reduction obtains for the second 'is' in *Kim is happier than Tim is* or in *I wonder where the concert is on Wednesday*, and countless other cases. The child needs no correction in *arriving* at this system: the very fact that 's is a clitic, a notion defined in advance of any

⁷ Chomsky affirms having once said that 'the child has a repertoire of concepts as part of its biological endowment and simply has to learn that a particular concept is realized in a particular way in the language' and adds that '[w]hen you read the huge Oxford English Dictionary ..., you may think that you are getting the definition of a word but you're not. All you are getting is a few hints and then your innate knowledge is filling in all the details and you end up knowing what the word means' (Chomsky 2000). Cook & Newson (2007) speak of a 'computational system' in the human mind which bridges meanings to sequences of sounds in one direction and sequences of sounds to meanings in the other. The lexicon is allegedly represented in the mind and the computational system relies on this mental lexicon.

experience, dictates that it may not occur in certain contexts. She now has a *reason* for the generalization that *is* may be pronounced as 's does not hold across the board⁸.

How this in fact goes on in the brain is of course far from clear. It is merely *assumed* that the brain is able to produce, store and evoke symbolic representations that inform or instruct the child when, in fact, symbolic representations can only be determined by conventions, not by neurons⁹. It seems also assumed that the child is able to understand a notion like *clitic*, for she is said to come to a conclusion about it. To reply that all the analysis, scanning, inferring and concluding she is said to do is done unconsciously or intuitively should not exempt the child from understanding, for what would be the sense of saying that the child arrives at a '*plausible analysis*' of something she does not, and *cannot*, understand? Moreover, this is all supposed to add to her antecedent *knowledge* of the language. But what 'knowledge' of 'clitics' can an average child 'already' have at 3? Again, adding 'unconscious' to 'knowledge' or replacing 'knowledge' with 'cognizance', as Chomsky was – precisely for such reasons – compelled to do, hardly allows him to retain all the epistemic activity – scanning, analysis, comparing, inferring and concluding – that allegedly goes on in the child's brain.

I suggest, then, that we replace this fantastical child-grammarians scenario with an explanation along the lines of: reducing the 's' in cases such as *Kim is happier than Tim is* makes the sentence incomprehensible; it gives: **Kim is happier than Tim's* – where the contraction misleadingly suggests possession, unlike such correct constructions as 'Sue's liking this'. Failure to communicate meaning may well be what discourages the child from reducing in such cases, besides the fact that she never hears it done in such contexts.

2. Principles Not Required

It is sometimes difficult to pick out what Chomsky is faulting, not only because his views have changed considerably over the years, but because of his lack of clarity. In speaking of language acquisition, Chomsky sometimes speaks of *language* not being acquired but at other times of *principles* not being acquired. So that when he claims children are seldom corrected by parents, one can easily disagree, but it is more difficult to disagree with the view that parents do not teach children linguistic *principles*, which leads Chomsky to conclude that these must be innate. For, on Chomsky's view – shared by nativists generally – to acquire a language, one must know the *principles* of language. Call this Chomsky's 'Principle Requirement'. The poverty of the stimulus argument crucially rests on this misleading requirement; that is, on the false assumption that the only way to *acquire and use* language in all its grammatical diversity is to know it in all its grammatical *complexity*. And as no child,

⁸ This paragraph is a faithful rendering of Anderson and Lightfoot (2000), 11.

⁹ For fleshed-out arguments on this, see Hutto & Myin (2013), Glock (2013) and Hacker (2007). Hacker: 'It is common among psychologists and cognitive neuroscientists to speak of internal representations in the brain. In so far as 'representation' signifies no more than a causal correlate in the brain of an external stimulus, this is innocuous. But it is evident that all too frequently it is meant to signify a *symbolic* representation. And it makes no sense to speak of semantic (symbolic) representations in the brain ... [f]or such representations are determined by conventions' (2007, 20-1).

or indeed adult, is ever taught all the grammatical complexity of language, we are forced to stipulate an innate universal grammar which endows the child with this complex knowledge – these principles – for how could she otherwise understand and speak her language?

But surely the fact that children know *language* by age three need not imply that children know *linguistic principles* by age three; what it does – more plausibly – imply is that knowledge of linguistic principles is not needed at all to know or use a language. The idea that children learn languages by discerning grammatical principles – or indeed, that this is the only possible way of acquiring a language – is an unnecessary and unsubstantiated stipulation which goes against all we experience and witness about how native languages are learned; namely, by directed and undirected repeated exposure to use, in a multiplicity of contexts. This basic observation is fleshed out in diverse ways by empirical studies.

Research shows that children systematize the language they hear based on the probability and frequency of forms, not on the basis of principles; they learn probabilistic patterns of word distributions, not syntactic rules; they generalize from cues, not from rules¹⁰. A series of experiments conducted by Hudson-Kam and Newport (2009) show that children ignore minor variations in linguistic input and reproduce only the most frequent forms – thereby regularizing and systematizing the inconsistent input and standardizing the language they hear around them. As Melodie Dye has also found:

Young children ... act like finely-tuned antennas, picking up the dominant frequency in their surroundings and ignoring the static. Because of this – because toddlers tend to pick up on what is common and consistent, while ignoring what is variable and unreliable – they end up homing in on and reproducing only the most frequent patterns in what they hear. In doing so they fail to learn many of the subtleties and idiosyncrasies present in adult speech (they will come to learn or invent those later). (2010)

It is precisely this picking up of the dominant frequency that explains characteristic errors such as children adding the suffix '-ed' to irregular verbs to mark the past, as in: 'He hitted me'¹¹. Such findings fit within a relatively new approach to language acquisition known

¹⁰ Studies in language development find that children use a wide variety of cues, including syntactic, semantic, and prosodic information, to learn language structure (Bates & MacWhinney, 1989). Bishop (2014): 'Current statistical learning accounts allow us to ... study the process of language learning. Instead of assuming that children start with knowledge of linguistic categories, categories are abstracted from statistical regularities in the input (see Special Issue 3, *Journal of Child Language* 2010, vol. 37). The units of analysis thus change as the child develops expertise. And, consistent with the earlier writings of Bates and MacWhinney (1989), children's language is facilitated by the presence of correlated cues in the input, e.g., prosodic and phonological cues in combination with semantic context. In sharp contrast to the idea that syntax is learned by a separate modular system divorced from other information, recent research emphasises that the young language learner uses different sources of information together. Modularity emerges as development proceeds.'

¹¹ Whereas the Chomskyan explanation here is that the regularity of such errors, and the fact that they are not based upon what the child hears, demonstrate that they are derived from the Universal Grammar. The child allegedly works through from the simplest possibilities offered by the UG to the more complex, until his own grammar is the same as the grammar of the mother-tongue. But even if we were to grant Chomsky the occurrence of such cerebral gymnastics, how does he explain that many children go on making mistakes of this kind into adulthood? I heard a man laughing at his companion who had just used the word 'sped' rather than 'speeded' (both are right), affirming that there's no such word. And how many of us are ever sure about when to use 'hung' or 'hanged'? It is bodies like the *Académie Française*, not UG, that legislate as to what is

as 'statistical learning'; as also 'similarity-based generalization' to explain how children are able to figure out how to use new words by generalizing about their use from similar words they already know how to use (Yarlett & Ramscar 2008). And of course, children also learn from repeated failure¹². So that word learning is a probabilistic, success and error-driven process, rather than a process of implicit rule-application.

Also, the nativist emphasis on the poverty of grammatical instruction and correction seems to rely on a restricted view of these. That children do not learn language with much in the way of *overt instruction*, as Fiona Cowie (2008) contends, is right if by this she means they are not taught linguistic principles or rules of syntax; but that they are often taught *language* is readily observable, as well as argued for by interactionist theorists of language acquisition, such as Lev Vygotsky's (with his notion of *collaborative learning*) and Jerome Bruner (with his *Language Acquisition Support System*)¹³. In fact, as Ted Schoneberger (2010) has shown, the crucial nativist claim that children find little linguistic reinforcement or corrective feedback is a myth. There is plenty of empirical evidence that children do encounter corrective data¹⁴. However, when presented with such evidence, Chomskians reject it on the grounds that whereas only *some* children are exposed to correct data, *all* children learn the correct rule. Indeed, a key justification for a Universal Grammar is that it explains why *all* children pick up the language correctly. As Green and Vervaecke write:

... the single most important datum to capture when modelling language learning is that children *virtually never* fail to learn language correctly, regardless of what kind of linguistic data they are exposed to early in life. (1997)

grammatically legitimate, and what changes are accepted, though it has a reactive rather than generative role – the evolution of language being mostly the spontaneous upshot of language users. As Ramscar, Dye & McCauley have found: 'children's overregularization errors both arise and resolve themselves as a consequence of the distribution of error in the linguistic environment, and ... far from presenting a logical puzzle for learning, they are inevitable consequences of it' (2013, 760).

¹² Ramscar & Yarlett (2007) and Ramscar, Dye & McCauley (2013) show the importance of expectation and error-driven learning processes in language acquisition. For example, when children erroneously expect an ungrammatical form that then never occurs, the repeated absence of fulfilment serves as a kind of implicit negative feedback which encourages them to correct their errors over time.

¹³ MacWhinney (1993) shows that language acquisition includes a 'rich armory of learning mechanisms' – including expressive and receptive monitoring, alongside competition, conservatism, complex, indirect and overt negative evidence, and cue construction – indicating that 'the logical problem of language learning is easily solved, and that there is really no logical problem of language acquisition at all'.

¹⁴ Moerk (1994) conducted a meta-analysis of 40 studies and found substantial evidence that corrections do indeed play a role, and that they are not only abundant but contingent on the mistakes of the child. Schoneberger (2010) cites findings that evidence (both positive and negative) available in the linguistic environment provides adequate constraints when learning a language. For example, children are provided positive evidence (a) when their grammatically correct utterances are directly reinforced by adults; (b) when their grammatically correct utterances are indirectly reinforced by adults by means of automatic reinforcement; and (c) when adults provide grammatically correct exemplars. Further, they are provided direct negative evidence when their grammatically incorrect utterances result in corrective feedback as well as indirect negative evidence by usually not being exposed to grammatically incorrect utterances. They also cite evidence to support the claim that reinforcement promotes language acquisition during naturally occurring parent-child verbal interactions.

I don't really know what to make of this; it sounds preposterous in that it is obvious that many children do fail to learn language correctly and that, in fact, many of them don't improve with age: a ride on the bus, a stroll in a shopping mall or even a university corridor or an hour of TV viewing will easily attest to the fact that many, if not most, people – children and grown children – do not speak grammatically correct English. For instance, we seem to be wiping out adverbs and are now doing 'amazing'. There is hardly a day when I don't hear some of the errors made in the following sentence: 'I'm feeling *more better* than I *done* yesterday though I'm not my *bestest*; but still, I should *of went* to the gym'. And other such common errors could be heard on the lips of many campaigning MPs, including Harriet Harman boasting about something 'we have *showed* [the country]' (2015 TV campaign appearance). Such errors are much too frequent and consistent to be dismissed as slips of the tongue. The point is not whether or not we ought to deplore this state of things – linguistic norms change¹⁵, like it or not – but that what we call 'grammatical' is not hard wired. Grammar is a mapping of language norms in use, not a map established in the brain in advance of all use, which linguists try to fathom.

Neither we, nor our brain, rely on or need to know any *principles* to form a correct sentence. If Universal Grammar (innate principles) were there to guarantee that the correct parameter is set, why do children (and adults) not eventually systematically produce the correct version of what they hear? Why, when she lives in an adult community which regularly produces expressions such as 'Me and her went to the gym' or 'They hurt theirselves', does the child use these mistaken expressions into adulthood? Or, alternatively correct herself if and when guided or exposed to correct usage. The number of adults who speak ungrammatically, and the fact that, as Dabrowska (2010) shows, agreement on well-formedness of complex sentences is far from universal in adults, easily belie the claim that, unless our language organs are severely impaired, we are universally either actual or potential perfect grammatical users of our native languages. What we speak is what we learn.

The poverty of stimulus argument took added impetus from findings by Derek Bickerton's examination of pidgins and creoles (1984). A pidgin has rudimentary grammar¹⁶; a creole is grammatically complex, but Bickerton found that the transition from one to the other is made within the space of one generation, and without the creole deriving its grammatical principles from the target language (that of the 'masters'), or from any of the substratum languages (one spoken by some subgroup of the labourers), suggesting that grammar is biologically hard-wired. Jean-Michel Fortis (2008) has shown Bickerton's evidence to be faulty. The complexification of Hawaiian pidgin seems to have been accomplished by speakers who were more or less passively bilingual, and therefore

¹⁵ There are of course many well-documented examples of this, such as changes in the grammar of Old English to that of Chaucer's Middle English.

¹⁶ Pidgins are basic or proto-languages developed as a means of communication by adults who do not share a common language. They are syntactically-impoverished languages, characterized by reduced syntax and vocabulary, no fixed order of words, with considerable variation from one speaker to another. However, a pidgin can evolve into a creole, which is a full-blown language. The argument here is that inasmuch as a full-blown language can be developed from an impoverished linguistic environment (with only vocabulary, but not grammatical principles drawn from a pidgin inasmuch as pidgins do not possess such principles in the first place), principles must be innate.

previously exposed to complex linguistic data; and new evidence shows that the complexification of pidgin to creole is not made within the space of one generation, but two¹⁷.

Another argument in support of nativism is the alleged 'domain specificity' of the language faculty: its independence from other aspects of cognition. This dissociation of language from the rest of our cognitive abilities is said to be evidenced by children who are linguistically prolific and yet present impaired intelligence; and, inversely, by cases where language ability is impaired in the presence of otherwise normal cognitive ability. The latter is known as 'Specific Language Impairment'. It is conceded, by Anderson and Lightfoot, that the homogeneity of cases that have been grouped under the SLI diagnosis is quite controversial, but this does not prevent them from using the argument. The controversy, I believe, is mostly to do with cases of feral children as the main evidence in support of SLI. The claim being that feral children fail to develop language despite normal cognitive abilities and exposure to language. But the objection that flares up here must be: that *normal cognitive abilities* are attributed to feral children in the first place. If not retarded from infancy and abandoned because of this, the psychological and physical abuse feral children suffered is alone sufficient to preclude normal language development¹⁸. The child can hardly be seen as growing up in a nurturing environment; the lack of language acquisition in later life may be due to the results of a generally abusive environment before – but also after – it was found, as the teaching methods used are often limited, if not repressive: the child is kept in (often harsh) experimental conditions, taught a native language by mere ostensive teaching of words rather than normal exposure and habituation in a form of life. It seems to me, then, that what such cases demonstrate is that language is *not* separable from cognitive development generally.

There is no space here to go into other kinds of case used to justify 'domain specificity', but as a general response, I would say two things: (1) there may be neural structures in the brain that are crucial to linguistic development, and the impairment of these may result in linguistic disabilities, but this neither justifies nor demands making those structures into a language acquisition device capable of storing and activating grammatical principles; (2) language acquisition is not separate from intelligence; it is one of our abilities that require

¹⁷ The transition occurring progressively. It is only in the second generation that pidgin is established, by speakers who have retained some of their native language. It is this stable and developed form of pidgin, constructed by the second generation, which gives birth to creole. The Hudson-Kam and Newport (2009) experiments mentioned earlier also suggest that creole languages do not support a universal grammar. In a pidgin situation (as also in the real-life situation of a deaf child whose parents were disfluent signers), children systematize the language they hear based on the probability and frequency of forms, and not on the basis of a universal grammar.

¹⁸ I cannot expand on this here but the literature is abundant – the cases of Victor and Genie being the most notable. In Genie's case, psychological and physical trauma was caused by her father who physically punished by her father if she made any sounds (Curtiss *et al* 1997, 127). As for Victor, he presented insensitivity to any feelings except joy and anger (e.g., he never cried; his eyes were without expression); he was virtually insensitive to noise and his sensitivity to temperature was different from the norm (e.g., he did not react to boiling water); he was unable to distinguish between a painting and an object in relief, and could not undertake mundane tasks like opening a door. See Singleton & Ryan (2004) against the validity of feral cases in support of the SLI hypothesis.

intelligence, but whose structural impairment does not necessarily impair our other intelligent abilities.

3. *Chomsky's Universal Grammar: a hopeless monster*¹⁹

The idea of a Chomskyan-style universal grammar has been contested on linguistic, neurobiological and anthropological fronts. Studies in descriptive linguistics have shown that there is no universal set of principles determining language; it is diversity, rather than universals, that is found at almost every level of linguistic organization. In their target article, which summarizes decades of cross-linguistic work, Evans and Levinson (2009) show just how few and unprofound the universal characteristics of language are²⁰, and how languages vary radically in sound, meaning, and syntactic organization. Where there are significant recurrent patterns in organization, these are better explained as stable engineering solutions satisfying multiple design constraints, reflecting both cultural-historical factors and the constraints of human cognition. So that '... the great variability in how languages organize their word-classes dilutes the plausibility of the innatist UG position' (2009, 429; 435). Research on syntax have brought Dunn *et al* to conclude that there is no universal set of rules determining the evolution of language; rather, it is cultural evolution that is the primary determinant of linguistic structure, at least with respect to word order (2011, 79).

The notion of an innate structure of mind imposing 'universals' has also been rejected from a biological perspective. As Wolfram Hinzen affirms, there is no clue, empirically, about the type of rule that would be able to organize neuronal connections that enable language competence; there is no biology specific to language; no Universal Grammar rooted in the human genome (2012a, 636). This segues with Christiansen and Chater's findings: 'a biologically determined UG is not evolutionarily viable' in that the processes of language change are much faster than processes of genetic change, so that language constitutes a 'moving target' both over time and across different human populations and cannot therefore provide a stable environment to which language genes could have adapted (2008, 489). And Terence Deacon abundantly demonstrates that Chomsky's scenario is unsupported by evolutionary anthropology which evidences a *gradual* adaptation of the human brain and vocal chords to the use of language rather than the sudden appearance of a language organ containing a complete set of parameters enabling all grammars²¹ – Chomsky's 'Big Bang' theory²².

¹⁹ A wink at what biologists call a 'hopeful monster' theory, described by Deacon as 'the evolutionary theorist's counterpart to divine intervention, in which a freak mutation just happens to produce a radically different and serendipitously better-equipped organism. The single most influential "hopeful monster" theory of human language evolution was offered by the linguist Noam Chomsky' (1997, 35).

²⁰ Views that espouse some universality don't need to appeal to nativism: Christiansen & Chater (2008) hold a non-formal conception of universals in which these emerge.

²¹ This is echoed by Christiansen & Chater (2008) whose research finds that it is non-linguistic constraints that have shaped language to the brain, and given rise to statistical tendencies in language structure and use. The question is not 'Why is the brain so well suited to learning language?', but 'Why is language so well suited to being learned by the brain?' Following Darwin, they argue that 'it is useful metaphorically to view languages as

'Don't think, but look!' (PI 66), writes Wittgenstein. By this, he means that we should draw our accounts of human life from *life*, from what we *do* rather than from presuppositions often generated by false analogies that lead to false requirements. And what we do shows that language acquisition is not due to the activation of principles but to the acquisition of a technique, a know-how, an ability. Children are not mini-grammarians. They don't need to 'cognize' and apply the Location Principle or any grammatical principle to acquire language. They use words in the same way we play tennis, without knowing what the mechanics are. Stipulating a Principle Requirement creates a problem where none was there, and appealing to innateness fabricates a solution where none is needed. Rather than 'language growing in us' – as Chomsky would have it – it is we who physiologically, psychologically and socially *grow* into language²³.

We can now move on to Wittgenstein, whose view of language development – in contrast to Chomsky's 'Big Bang' theory – seems to me ontogenetically and phylogenetically consistent and coherent.

4. Wittgenstein's Social Account of Language Acquisition

4.1 The Primitivity of Action: the deed, not the word

'In the beginning was the deed'
Wittgenstein, quoting Goethe (OC 402)

One of the important things Wittgenstein said about language is that it has its root in gesture – or, as he also put it, in 'action', and more precisely: 'reaction' or 'instinct': 'What we call meaning must be connected with the primitive language of gestures' (BT 24). By this, he means instinctive gestures and reactions which – first through evolution, then through enculturation – get replaced by words. This – Wittgenstein's 'primitivism' (Canfield 1997, 258) – prompted Michael Tomasello to realize that '[i]f we want to understand human communication, ... we cannot begin with language'; contrary to primatologist dogma, apes' *gestures*, not their vocalizations, are the precursors of human language (2008, 59; 53-5)²⁴.

'organisms' – that is, highly complex systems of interconnected constraints – that have evolved in a symbiotic relationship with humans' (2008, 490).

²² This view of the sudden appearance of language as a kind of evolutionary accident where humans, to the exclusion of all other animals, were somehow accidentally blessed with a fully functioning prefabricated language organ (see Chomsky 1988) has been found, Green & Vervaecke (1997) concede, hardly plausible. However, they retort: 'ironically, ... the *real* Big Bang theory is, as far as we now know, true! A substantial critique of the implausibility of 'catastrophic' or 'big bang' theories of brain evolution to account for humans' unique linguistic capacity can be found in Deacon (1997).

²³ Chomsky is averse to saying that language acquisition is a learning at all; it is, to him, more akin to *growing* than to learning: 'In certain fundamental respects we do not really learn language; rather grammar grows in the mind' (1980a, 134); 'language development really ought to be called language growth because the language organ grows like any other body organ' (1983). Cook & Newson: 'Acquisition of language is, to Chomsky, learning in a peculiar sense: ... it is not like learning to ride a bicycle, where practice develops and adapts existing skills. Instead it is internal development in response to vital, but comparatively trivial, experience from outside' (2007, 185).

²⁴ For an *aperçu* of Wittgenstein's impact in the field of language acquisition, see Nelson (2009).

When Wittgenstein writes that '[t]he study of language games is the study of primitive forms of language or languages' (BB 17), he does not, by this, mean words or symbols, but reactions:

The origin and the primitive form of the language game is a reaction; only from this can more complicated forms develop.

Language – I want to say – is a refinement. 'In the beginning was the deed.'
(CE 395)

Language, then, is a refinement, or 'an extension', of our primitive behaviour²⁵; it emerges from the *development* of some of our animal or natural reactions. Not just *any* natural reaction – not singular or idiosyncratic ones, like tics – but our *shared* natural reactions; what Wittgenstein calls 'the common behaviour of mankind' (PI 206): reactions such as crying when in pain or sad; smiling when glad; jumping when startled; gasping or screaming when afraid; but also reacting to someone's suffering. He writes:

In its most primitive form [the language-game] is a reaction to somebody's cries and gestures, a reaction of sympathy or something of the sort. (CE 414)

These instinctive common reactions or action patterns are, Wittgenstein says, the *prototypes* of our modes of thought (RPP I, 916); of our concepts. And so, the basis for the development of language is constituted by a number of such distinct instinctive, behavioural patterns which John Canfield calls 'proto-language games' (1996, 128). Without these behavioural patterns, there would be no language²⁶. This is the case phylogenetically as well as ontogenetically; for these natural configurations of behaviour – such as: '[t]he natural, untutored behaviour of one pre-linguistic hominid helping another it sees is hurt' – are part of the species' inheritance' (*ibid.*). So that ontogeny recapitulates phylogeny.

4.2 Training

In the *Investigations*, Wittgenstein writes that the primitive forms of language are those used by the child when it is learning to talk and that, here, 'the teaching of language is not explanation, but training' (PI 5). Why? For an obvious reason: in the learning of a first language, the initiate has only instinct and reactions but no language at her disposal, and so

²⁵ 'Being sure that someone is in pain, doubting whether he is, and so on, are so many natural, instinctive kinds of behaviour towards other human beings, and our language is merely an auxiliary to, and further extension of, this relation. Our language-game is an extension of primitive behaviour.' (Z 545). 'What, however, is the word "primitive" meant to say here? Presumably, that the mode of behaviour is *pre-linguistic*: that a language-game is based *on it*: that it is the prototype of a mode of thought and not the result of thought.' (RPP I, 916). Wittgenstein speaks of *primitive* or *animal* behaviour in the phylogenetic as well as the ontogenetic sense. Here is an illustration of the phylogenetic primitivity of our concepts: '(An ape who tears apart a cigarette, for example. We don't see an intelligent dog do such things. The mere act of turning an object all around and looking it over is a primitive root of doubt' (RPP II, 345).

²⁶ Wittgenstein: 'it is characteristic of our language that the foundation on which it grows consists in *steady ways of living, regular ways of acting*' (CE 397; my emphasis). Our acquiring concepts, such as pain, requires that we have appropriate (i.e. normal) human reactions: 'If a child looked radiant when it was hurt, and shrieked for no apparent reason, one couldn't teach him to use the word "pain"' (LPP 37).

the learning of a native language will *have* to do with action or behaviour – language at first playing only a background music role. This is why language cannot take its impetus from explanation: 'Language did not emerge from some kind of ratiocination' (OC 475).

Language, then, is an extension of our patterned non-linguistic behaviour through training. On Wittgenstein's view, this should include at least one competent trainer – that is, a reasonably adept user of the language, endowed with enough pedagogic ability to mould or shape the child's responses to the training so that they end up in harmony with the norm²⁷. Of course, by 'training' here, is not meant anything formal but the kind of repeated direct and indirect guidance that is effected through intersubjective interaction in various contexts. More formal teaching may – it does not always – come later²⁸. But it is clear that for Wittgenstein, language acquisition is not a one-track affair: he rejects the Augustinian view of it as resulting from a mere nominalization process, and describes it in terms of habituation and enculturation.

So how is the child trained to go from proto-language games to language, from her instinctive reactions to language? '... how, asks Wittgenstein, is the connexion between the name and the thing set up?' – and he replies:

This question is the same as: how does a human being learn the meaning of the names of sensations? – of the word 'pain' for example. Here is one possibility: words are connected with the primitive, the natural, expressions of the sensation and used in their place. A child has hurt himself and he cries; and then adults talk to him and teach him exclamations and, later, sentences. They teach the child new pain-behaviour. (PI 244)

So the *connection* between the name and the thing is not made by an act of ostension, not by merely hooking gestures on to their public referents, but by processes of drill or habituation that are similar to stimulus-response conditioning, but that must be supplemented by training *into the practice* in which those words are used²⁹. Note: what the child is taught in learning to replace his primitive reactions with words is new *behaviour*; that is, in first picking up the linguistic expression, the child is not *describing* with it or *referring* with it, but still *reacting* with it.

Ostensive *teaching*, as opposed to mere ostensive *definition*, involves behavioural conditioning: the child is taught, through repetition and exercises, to utter certain words in certain contexts or situations. These drills are used to tap and channel the child's natural reactions. What we witness in these initial stages is not yet language, but 'processes

²⁷ On Philippe Narboux' view, training is a necessary but insufficient condition for the learning of a *native* language whereas *second-language* acquisition doesn't require it, and can rely on nothing other than ostensive definition because it relies on previous training (2004, 136).

²⁸ Once the child has some language, there will be more explanatory teaching, and perhaps the odd transmission of some linguistic principles (though not usually of the 'clitic' sort). Wittgenstein talks about teaching as well as training: (e.g. Z 318 & 186).

²⁹ Cf. Medina (2002), 173. As psychologist Derek Montgomery also observes, if the carer repeatedly uses the verb 'want' while interpreting the infant's behaviour in certain contexts, it is 'reasonable to suspect that when the verb emerges in the child's lexicon it will be in familiar contexts such as [those] where the child has repeatedly heard it being used. The meaning of the term, like the meaning of the prelinguistic gesturing, is bound up in the role it plays within such contexts' (2002, 372).

resembling language' (PI 7); for a language is not the mere repetition of certain sounds in certain contexts and after certain prompts. Wittgenstein is not a behaviourist. Drill is not enough; beyond mere conditioning (cf. PI 6), a normative *attitude* towards utterances, towards how things are to be done, must be inculcated in the child, so that it can learn to regulate itself³⁰. And acquiring a normative attitude demands nothing less than being enculturated. It is thanks to her acquiring this normative attitude that the child is eventually able to go on, on her own; to proceed *from other-regulation to self-regulation* (Medina 2002, 165). Successful enculturation means the child can then judge for herself that in a particular instance a word or phrase makes sense, not by comparing it to a benchmark, context-free, use but on the basis of her experience of multiple language-games in which the word or phrase is employed. For Wittgenstein, the acquisition of language is the acquisition of a technique, skill or capacity, and capacities are flexible to individual and occasion; they allow for – and indeed, are basic to – productivity and creativity.

5. Technique: Wittgenstein's answer to the productivity problem

To understand a language means to be master of a technique.

Wittgenstein (PI 199)

So that 'technique', I suggest, is Wittgenstein's answer to the nativist problem of productivity, which is how to account for our capacity to produce an unlimited number of novel utterances from the limited set of grammatical rules and the finite set of terms we acquire from experience. On Wittgenstein's view, we are able to extend our limited acquired knowledge of language to new situations and contexts precisely because the teaching of language is not a teaching of principles but the transmission of a technique, which does not aim for total regulation, but for self-regulation.

It is criteria that determine whether a speaker is following a rule or using a word in accordance with the norm that is being inculcated. These criteria are public, not private; they can be transmitted to the child and invoked to guide and correct him in his attempts to use the words he is being taught. The child's various attempts are guided (encouraged/discouraged) until enough training allows her to grasp what *sorts of contexts* are propitious for the use of the word: semantic development involves precisely 'becoming increasingly sensitive to how characteristics of different contexts constrain the words one can use' (Montgomery 2002, 373)³¹. However, though constraint is necessary, there is no exhaustive determination of use

³⁰ 'Our children are not only given practice in calculation but are also trained to adopt a particular attitude towards a mistake in calculating [variant: '... towards a departure from the norm']' (RFM VII 61, p. 425) – that is, children are habituated into standards of correctness of the practice in question, and thereby formed to act and react in particular ways; they are thus trained to master a technique.

³¹ For Chomsky, in contrast, our words are informed by the brain; they get their meaning from internal meanings (which are abstract mental representations), and it is the brain that communicates meaning: the human mind bridges the gap between external sounds and internal meanings (which are abstract mental representations) via a 'computational system' that relates meanings to sequences of sounds in one direction and sequences of sounds to meanings in the mind in the other. The mind changes the representation of language used by the computational system into the general concepts used by the mind, called 'the conceptual-intentional system', i.e. *moon* is

but an indication of proper use (the use is *constrained, not shackled*), which allows for and explains creativity/productivity.

We might make an analogy here with a dog that is trained not to bite: the dog will not only not bite the people present at the training, but not bite in all similar contexts (e.g. unthreatening contexts). Or again, when the child is taught to open a door, she doesn't just learn to open *that* white, single panelled door which her mother is using to teach her, but all doors that she will come across in experience – whether they be white, black, double-panelled, glass and so on. This is where the teaching of a technique surpasses, say, ostensive definition. As Wittgenstein writes:

Teaching which is not meant to apply to anything but the examples given is different from that which '*points beyond*' them.³² (PI 208)

To suggest that we need a Language Acquisition Device to explain how we can produce an unlimited number of new and correct sentences from limited input is like suggesting we need a Manoeuvre Acquisition Device to explain the potentially unlimited variety of manoeuvres that a cyclist is able to execute from the basic or limited training he has received (cf. Rorty 2004, 225). Wittgenstein's answer to the productivity problem is encapsulated in this passage: 'Yes, there is the great thing about language – that we can do what we haven't learnt' (LPP 28).

So that, *contra* Chomsky, first-language acquisition is *essentially* social; it requires that at least one member of the linguistic community mould the child's primitive reactions and proto-language games into language-games, bringing the child, through a process of enculturation, to assimilate, conform to and apply the standards of correctness of its linguistic community. Acquiring language is like learning to walk: the child is stepped into language by an initiator and, after much hesitation and repeated faltering, with time, multifarious practice and repeated exposure, it disengages itself from its teacher's hold and is able, as it were, to run with the language.

6 Training is not enough

This may be a good time to bring up Chomsky's argument that while human babies and animals can both be trained, if they are exposed to exactly the same linguistic data, the human

connected to the concept of 'earth's satellite'. Going in the opposite direction, while speaking the mind has to convert the concepts into linguistic representation for the computational system, i.e. 'earth's satellite' is converted into *moon*. (Cook & Newson 2007, 6). In contrast to this mentalist view, for Wittgenstein (echoed here by Montgomery), it is in social practices that the meaning of words and the standards for their use are established. Meaning, as Wittgenstein says, is 'in use' – out there – not in the head, not in some mental repository.

³² Wittgenstein's rule-following argument shows precisely that generating new sentences is nothing but an instance of knowing how to go on, 'how to extend the speech that [we] have into new contexts' (Bruner 1983, 39). As H.-J. Glock notes, the early Wittgenstein's was also concerned with what is now known as the problem of 'the creativity of language': the number of propositions being indefinite although the number of words is finite (NL 98; TLP 4.02, 4.027 etc.) (1996, 298).

child will acquire language, but the kitten will not. For Chomsky, to suggest that language is *not* innate is to imply that no crucial and relevant internal nature differentiates a child from a rock or a rabbit; so that if we put a rock, a rabbit and a child in an English-speaking community, they should all be able to learn English (Chomsky 2000). It is in order to avoid coming to this conclusion, he says, that we must accept that the relevant capacity the child has which the stone and rabbit lack is the 'language acquisition device' (LAD). But this, I suggest, is like using a hammer to crack an egg. We hardly need resort to a language acquisition device to mark a difference here – internal or otherwise – between humans and animals (to say nothing of rocks). Before resorting to something as *ad hoc* as a language organ to account for the difference, we might, for a start, evoke obvious, traceable physiological requirements such as the anatomical structures necessary for the vocalization of human speech that are lacking in rabbits and rocks. Or indeed as Elizabeth Bates does, to differences between human brains and, say, those of chimpanzees, but differences – she insists – that do not require us to postulate a language organ (1993, 8). However, the difference is not only physiological³³.

Rebecca Saxe, a cognitive scientist at MIT, argues that the key difference between apes and humans seems to be that we have *explicit teaching* while animals have only imitation. Chimps may use sticks as rudimentary tools, but they learn such skills through observation and mimicry as well as trial and error, rather than direct instruction. Humans learn through all of these, but teaching may be the signature skill of our species. Also, a part of teaching is 'triadic attention': being able to work out the coordination of my attention and the other person's attention on this third thing: the task I'm trying to teach. I have to pay attention to all three elements continuously. Finally, it seems like 'it's not just a cognitive capacity that's necessary for teaching; there's this other thing, which is *wanting to teach*. That may be even more critical. We need to understand that somebody else is unable to do what we're doing, and also have some reason, motivation, desire to help that person learn it. As Saxe says: 'That desire to teach seems to be really pervasive in humans and may be mysteriously missing in apes' (2008). Humans have both the passion and the mental skill to teach each other.

So that the inability for rabbits and stones to speak and teach goes some way in explaining why they don't acquire language and we do. But also the fact that children *expect* (and like) to be taught makes a vital difference, and this is visible in their capacity for attention. Though the learner's *triadic* attention is a development, attention for teaching appears before language kicks in. Apes can't get into that; although they do things together and coordinate their actions together, they have no *shared goals based on shared commitment*. Kids, on the other hand, have these naturally, almost immediately.

Another distinguishing feature of humans is what John Haugeland calls 'norm-hungriness' (2002, 22) – an expectation of norms that parents try to teach us. Haugeland

³³ Actually, the deeper confusion here is that unless we assume that language is innate, exposure to English would have to result in the child, the rabbit and the rock learning **English**. As Witt said: 'If a lion could talk, we wouldn't be able to understand it' (PPF 327), for he wouldn't – couldn't – speak 'human'. Learning human speak (e.g., English) – takes enculturation in a human form of life, and that presupposes shared behavioural reactions and responses. I won't even bother about the rock ...

claims the fundamental divide between humans and animals lies in our essential use of norms. Humans display a norm-susceptibility, and a deep 'norm-hungriness'; deep, both in that the desire is almost insatiable, and in that the norms are unique in their richness and complexity (*ibid.*, 31).

7. Wittgenstein's universal grammar

I'd like to conclude this paper by suggesting that we replace Chomsky's universal grammar with Wittgensteinian grammar. Inasmuch as the latter is all the grammar we require to possess a language, and inasmuch as its acquisition can only be achieved through enculturation, it makes an otiose monster out of Chomsky's hopeful monster.

As linguist Dan Everett (2012) says: 'universal grammar doesn't seem to work, there doesn't seem to be much evidence for that'; but what, he asks, can we put in its place? His answer is a Wittgensteinian one: 'A complex interplay of factors, of which culture, the values human beings share, plays a major role in structuring the way that we talk and the things that we talk about'. In view of the implausibility of Chomsky's UG³⁴, Wolfram Hinzen seeks to rehabilitate universal grammar as internally linked to human nature, without that centrally involving notions of innateness, essence or modularity (2012b, 55). Gone is the idea of grammar as an autonomous module described by arbitrary formal rules; grammatical theory is a theory of the mode of thought that we find expressed in human language and that is manifest in human culture (2012a, 643). Hinzen's view of grammar seems apt to resonate with Wittgenstein's in that Wittgensteinian grammar is a description of the rules or bounds of sense that determine, and are embedded in, our use of language as a result of the evolution of languages in the various cultures of human life.

For Wittgenstein, there is no question that language depends on grammar; but Wittgensteinian grammar is a different animal from Chomsky's. Whereas Chomsky construes grammar as a set of arbitrary *principles* existing in advance of use or practice, for Wittgenstein, it is not principles, but rules or norms that are necessary to the existence of a language, and these do not pre-exist language but are inextricably bound up with its practice. As John Canfield puts it: 'The practice underlies the rule rather than vice-versa' (1975, 114), which is to say that grammatical rules merely express or *bring out* our *normative* use of words and expressions. In fact, rules of grammar, as Wittgenstein conceives them, are nothing like linguistic principles; they are garden-variety or ordinary expressions (reminders) of the norms that have been regulating our meaningful use of words (e.g., 'A bachelor is an unmarried man'; 'This \emptyset is what we call a table'; 'Red is darker than pink'; 'A rod has a length'). The child assimilates these norms as it assimilates the language – through guidance (which may, but need not, involve explicit reminders of the norms) in, and exposure to, correct usage.

Learning the meaning of a word is nothing but learning how it is used; that is, assimilating the *norms that govern its use* – what Wittgenstein calls its *grammar*. We could

³⁴ Hinzen: '... controversies about UG abound and the enterprise is widely rejected as ill-conceived and unfounded' (2012a, 335n).

not speak of the number of people in a room if we did not have a norm- or rule-governed use of the words 'people', 'number' and 'room'. So that if I were to say that I have counted 30 people in the room, of whom two dogs and a cat, I would not be using the words 'people', 'dog' and 'cat' correctly. I would not be speaking grammatically. Grammatical norms determine what it makes sense for us to say also by excluding certain combinations of words from meaningful use³⁵.

The grammatical rules I have been describing are what we might call 'thin' and 'local': they are norms for the use of words, expressions, and numbers³⁶ ('2+2=4' is for Wittgenstein as much a rule of grammar as 'A rod has a length'). But Wittgenstein came to realize that many of our grammatical rules are 'thick' or 'reality-soaked'³⁷: they are *conditioned* by empirical reality. So, for example: 'The earth has existed for a long time'; 'Human beings have bodies'; 'If someone's head is cut off, he is dead and will never live again'; 'Babies cannot look after themselves'; 'Human beings are born, grow old and die; they need air, water and sleep; they usually smile or laugh when they're happy; cry when they're sad or in pain'.

Now of course Wittgenstein is well aware that these look like run-of-the mill empirical propositions, not expressions of rules of grammar. Yet he notices that these are propositions about which 'no doubt can exist if making judgments is to be possible at all', and so they do not, in fact, *function* as empirical propositions³⁸ but as *logical* propositions: they 'form the foundation of all operating with thoughts (with language)' (OC 401). Failing, in our language-games, to 'assume' the indubitability of such propositions would result in a failure to make sense: for an English speaker to ask, in a non-sci-fi context, whether 'human beings have bodies' would be as nonsensical as asking whether a 'rod has a length'. Either question would make us suspect, first, the speaker's proficiency in English and if that were ascertained, her sanity. Inasmuch, then, as our language-games are *logically* hinged on such putative 'propositions', Wittgenstein takes them to have a *grammatical* role ('logical' and 'grammatical' are synonymous for the later Wittgenstein). Such 'propositions' are in fact expressions of rules of grammar.

Their seemingly empirical nature is due to these rules of grammar being *conditioned* by 'very general facts of nature' appertaining to 'the natural history of human beings'³⁹ (PI 230,

³⁵ Wittgenstein does not think his conception of grammar contrasts with the grammarian's; he insists that 'any explanation of the use of language' is 'grammar' (MWL 69). On his view, 'A rod has no length' is as ungrammatical as 'A rod length has'; but as he concedes (to Moore), the former violation of grammar is of interest only to the philosopher (*ibid.*), whereas syntax – albeit also part of grammar – is not the part philosophers are interested in. Wittgenstein leaves it to grammarians to bring out the syntactic aspect of use. Grammarians and philosophers may find it of interest to map grammatical rules, but this does not make the apprehension of rules as such relevant to language acquisition. In picking up the correct syntactico-semantic use of language – its grammar – the child is not picking up rules as such, but simply, to repeat: correct use.

³⁶ I am generalising for simplicity's sake. For a more nuanced view of Wittgenstein's conception of grammatical rules, see Glock (1996), 150-55, and Moyal-Sharrock (2004).

³⁷ I owe this expression to Bernard Harrison (1991, 58).

³⁸ That is, we are interested in the fact that about certain empirical propositions no doubt can exist if making judgments is to be possible at all. Or again: I am inclined to believe that not everything that has the form of an empirical proposition *is* one.' (OC 308)

³⁹ Note: *conditioned*, not *justified* by facts or *inferred* from them. This is what precludes their being empirical propositions. Rules are not empirically or epistemically grounded in reality, though they may be 'caused' by reality (OC 131, 429, 474). This is why Wittgenstein writes: 'The rule we lay down is the one most strongly

415): 'it is characteristic of our language that the foundation on which it grows consists in steady ways of living, regular ways of acting' (CE 397). Their being conditioned by facts that unassailably pertain to the human form of life makes them *universal* rules of grammar – that is, laws of thought belonging to the 'scaffolding of thought' of any normal human being. They are the bounds of sense from which *any* human being must begin to make sense⁴⁰. Such rules, I suggest, constitute alongside our local grammars, a universal grammar⁴¹.

8. Conclusion

We have seen that Chomsky's UG faces multiple objections from multiple fronts – objections it cannot answer. This, however, has not managed to take the wind out of UG's sails⁴². Myths are difficult to debunk. For, if anything transpires from even a glance at Chomsky's position, it's that it does have the trappings of myth. On the other hand, Wittgenstein's conception of language acquisition – resulting as it does from observations so basic as to be incontestable – is of the order of 'perspicuous presentation'; it merely makes obvious the *familiar unseen*: 'We want to understand something that is already in plain view. For this is what we seem in some sense not to understand' (PI 89).

There is nothing mysterious, speculative or *ad hoc* about Wittgenstein's view of language acquisition: no positing of hidden principles lodged in the brain and requiring 'cognisance'; all that is needed is the (plain-to-see) ability to be initiated in and master the technique of language through enculturation. Any more detail than this must be left to

suggested by the facts of experience' (AWL 84). For clarification regarding the nonempirical nature of 'thick' rules of grammar, see Moyal-Sharrock (2013).

⁴⁰ Some people may hold beliefs that seem to violate – and so could not condition – universal rules of grammar. For example, in the Trobriand Islands, some women, called *Yoyova* or flying witches, are believed to have the capacity to fly. It is, however, also (accommodatingly) believed that they either leave their bodies behind when they do this, or have doubles in the form of fire-flies etc. do it for them (Young, 207). The universal rule of grammar: 'Human beings cannot fly unaided' is therefore not actually transgressed. Any attempt to ignore or transgress it *in action* – such as a *yoyova* attempting to *actually* fly off a cliff (without 'leaving her body behind') – must be seen as pathological. For any local belief that seems to violate a universal rules of grammar, such accommodating measures will always be found. There is no *normal* transgression of a universal rule. To genuinely think or act on the basis of such rules of thought as – 'I can fly unaided' or 'Only I exist' is a pathological problem, not a doxastic option. See Moyal-Sharrock (2007) for a discussion of *local* and *universal* rules of grammar drawn from *On Certainty*.

⁴¹ They are conditioned by 'extremely general facts of nature – such facts as are hardly ever mentioned because of their great generality' (PI, p. 56) – including the 'common behaviour of mankind' – behaviours such as breathing, eating, walking, hoping, dying, speaking, thinking, giving orders, asking questions, telling stories, having a chat. It is this common behaviour that constitutes the universal 'system of reference' which conditions what might be called, though in obvious contrast to Chomsky, the 'universal grammar' of mankind – that grammar by means of which any human being can understand a foreign language (PI 206).

⁴² As Vyvyan Evans testifies, Chomsky's views are 'established fact in many of the linguistics textbooks currently in use in many of the stellar universities throughout the English-speaking world. I was trained using these textbooks, and they are still compulsory reading for today's undergraduate and graduate students – tomorrow's researchers, educators and language professionals – even at the university where I teach and work. University students are regularly told that there is a Universal Grammar, that language is innate, that language is incommensurable with non-human communication systems, and that all languages are essentially English-like' (2014, 19-20).

empirical research, but of the non-mythopeic kind⁴³ – the kind, instead, that Ramscar and Yarlett are gesturing at here:

Language is ultimately a cultural capacity (Tomasello, 1999; Wittgenstein, 1953 [PI]); arguably, it is the capacity for culture that sets *Homo sapiens* apart from our closest neighbors. Understanding how the processes of imitation that appear to be key to the acquisition and establishment of cultural common ground interact with the processes that allow humans to exert more cognitive control over their responses, and thus achieve agency across the course of cognitive development, may ultimately result in a much deeper understanding of our capacity for, and the nature of, both language and culture. (2007, 952)

Constancy of meaning, linguistic communication and expression are made possible by a grammar, but this grammar is not a set of innate linguistic principles biologically programmed in our brains in advance of all use, and enabling the inner growth of language. In contrast to Chomsky's grammar, there is nothing inexplicable about Wittgenstein's grammar. On his view, human grammar is internally related to human life and action – this is what his concept of the *language-game* conveys, as well as his often-reiterated conviction that 'at the beginning is the deed'. If universality is the motivating force behind nativism, there is in our shared natural history and in our shared instinctive reactions, universality enough. Our grammars are embedded in our use of language and conditioned by our forms of life; it is therefore in the 'stream of life' – and not in the human brain – that Wittgenstein rightly locates these grammars.

⁴³ That it is necessary to have a well-functioning brain to achieve language acquisition does not make the brain the locus of language acquisition – or part of it, a 'language organ'. The brain is a mere mechanical enabler: its proper functioning is necessary to our acquiring and using language in the same way it is necessary to our ability to walk or digest – without it implying the existence of walking and digestion organs in the brain.

References

- Anderson, S. R. & Lightfoot, D. W. (2000) 'The Human Language Faculty as an Organ' *Annu. Rev. Physiol.* 62, 1-23.
- Bates, E. (1993) 'Comprehension and Production in Early Language Development: Comments on Savage-Rumbaugh et al.' in *Monographs of the Society for Research in Child Development* 58: 3-4, 222-42.
- Bates, E. and MacWhinney, B. (1989) Functionalism and the competition model. In B. MacWhinney and E. Bates (eds), *The Crosslinguistic Study of Sentence Processing* Cambridge: Cambridge University Press, 3-76).
- Bickerton, D. (1984) 'The language bio-program hypothesis' *Behavioral and Brain Sciences* 7:173-212.
- Bishop, D. (2014) What Chomsky doesn't get about child language. Figshare.
<http://dx.doi.org/10.6084/m9.figshare.1030403>
- Bruner, J. (1983) *Child's Talk: Learning to use language*. New York: Norton.
- Canfield, J. V.
1975 'Anthropological Science Fiction and Logical Necessity' in *The Philosophy of Wittgenstein*; Vol. 10: *Logical Necessity and Rules*. Ed. J. V. Canfield. NY: Garland, 1986, 105-117.
1996 'The Community View' *The Philosophical Review* 105: 4, 469-88.
1997 'Wittgenstein's Later Philosophy' in *Philosophy of Meaning, Knowledge and Value in the Twentieth Century*. *Routledge History of Philosophy*, Vol. X. Ed. J. V. Canfield. London: Routledge, 247-85.
- Chomsky, N.
1972 *Language and Mind*. Extended edition. Harcourt, Brace & World.
1980a *Rules and Representations*. Oxford: Blackwell.
1980b 'On cognitive structures and their development: a reply to Piaget. In *Language and Learning: The Debate between Jean Piaget and Noam Chomsky*. Ed. M. Piatelli-Palmarin. Cambridge: Harvard University Press, 1980, 35-54.
1983 'Things No Amount of Learning Can Teach' Noam Chomsky interviewed by John Gliedman. *Omni*, 6:11. <http://www.chomsky.info/interviews/198311--.htm>
1988 'Language and the problems of knowledge' The Managna Lectures. MIT Press.
1993 *Language and Thought*. London: Moyer Bell.
1995 *The Minimalist Program* (MIT Press, 1995)
2000 'The "Chomskyan Era"', Extracted from *The Architecture of Language*, 2000. Online.
- Christiansen, M. H. & Chater, N. (2008) 'Language as shaped by the brain' *Behavioral & Brain Sciences* 31, 489-58.
- Cook, V.J. & Newson, M. (2007) *Chomsky's Universal Grammar: An Introduction*. Third Edition. Oxford, UK: Blackwell Publishing.
- Cowie, F. (2008) 'Innateness and Language'. *Stanford Encyclopaedia of Philosophy*:
<http://plato.stanford.edu/entries/innateness-language/#toc>
- Curtiss, S, et al. (1974) 'The development of language in Genie: a case of language acquisition beyond the "critical period"'. *Brain and Language* 1:1, 81-107.
- Dabrowska, E. (2010) Native vs. expert intuitions: an empirical study of acceptability judgements. *The Linguistic Review*, 27, 1-23.
- Deacon, T. (1997) *The Symbolic Species: the co-evolution of language and the human brain*. Penguin.
- Dunn, M., Greenhill, S. J., Levinson, S. C., & Gray, R. D. (2011). Evolved structure of language shows lineage-specific trends in word-order universals. *Nature*, 473, 79-82.

Dye, Melody (2010) 'The Advantages of Being Helpless: Human brains are slow to develop – a secret, perhaps, of our success' *Scientific American*, February 2010.
<http://www.scientificamerican.com/article/advantages-of-helpless/>

Evans, N. & Levinson, S. C. (2009) 'The myth of language universals: Language diversity and its importance for cognitive science' *Behavioral and Brain Sciences* (2009) 32, 429-92.

Evans, V. (2014) *The Language Myth: Why language is not an instinct*. Cambridge: Cambridge University Press.

Everett, D. 'There is no such thing as universal grammar' Interview by R. McCrum. *The Guardian* 25 March 2012 <http://www.theguardian.com/technology/2012/mar/25/daniel-everett-human-language-piraha>

Fortis, J.-M. (2008) Le langage est-il un instinct? Une critique du nativisme linguistique de Chomsky à Pinker. *Texto!* 13:4, <http://www.revue-texto.net/index.php?id=1870>.

Glock, H.-J.
 ---- (1996) *A Wittgenstein Dictionary*. Oxford: Blackwell.
 ---- (2013) 'Animal Minds: A Non-Representationalist Approach' *APQ* 50:3, 213-32.

Green, C. D. & Vervaecke, J. (1997) But What Have You Done for Us Lately? Some Recent Perspectives on Linguistic Nativism. In Johnson, D.M. & Erneling, C.E., eds *The Future of the Cognitive Revolution* (OUP, 1997), 149-63. Online (not paginated).

Hacker, P. M. S. (2007) 'The Relevance of Wittgenstein's Philosophy of Psychology to the Psychological Sciences' forthcoming in *Proceedings of the Leipzig Conference on Wittgenstein and the Psychological Sciences*.

Harrison, B. (1991) *Inconvenient Fictions: Literature and the Limits of Theory*. Yale University Press.

Haugeland, J. (2002) 'Andy Clark on Cognitive and Representation' in *Philosophy of Mental Representation*, ed. H. Clapin. (Oxford: Clarendon Press, 2002), 21-36.

Hinzen, W.
 ---- (2012a) 'The philosophical significance of Universal Grammar' *Language Sciences* 34 (5): 635–49.
 ---- (2012b) 'Human Nature and Grammar' *Royal Institute of Philosophy Supplement* 70, 53-82.

Hudson Kam, C.L., & Newport, E.L. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology*, 59, 30-66.

Hutto, D. D. & Myin, E. (2013) *Radicalizing Enactivism: Basic Minds without Content*. Cambridge, MA: MIT Press.

MacWhinney, B. (1993). The (il)logical problem of language acquisition. In Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society. Hillsdale, NJ: Lawrence Erlbaum Associates.

MacWhinney, B. *et al* (1989) 'Language Learning: Cues or Rules' *Journal of Memory and Language* 28, 255-77.

Medina, J. 2002 *The Unity of Wittgenstein's Philosophy: Necessity, Intelligibility, and Normativity*. Albany: SUNY Press.

Moerk, E. L. (1994). Corrections in first language acquisition: Theoretical controversies and factual evidence. *International Journal of Psycholinguistics*, 10, 33-58.

Montgomery, D. E. (2002) 'Mental Verbs and Semantic Development' *Journal of Cognition and Development*, 3:4, 357-84.

Moyal-Sharrock, D.

---- (2004) *On Certainty* and the Grammaticalisation of Experience in *The Third Wittgenstein: the post-Investigations works* (Aldershot, UK: Ashgate, 2004), 43-62.
 ---- (2007) *Understanding Wittgenstein's On Certainty*. Basingstoke: Palgrave Macmillan.
 ---- (2013) Realism, but not empiricism: Wittgenstein versus Searle in *A Wittgensteinian Perspective on the Use of Conceptual Analysis in Psychology*. Eds T. P. Racine and K. L. Slaney (Palgrave Macmillan, 2013), 153-71.
 ---- (Forthcoming) Wittgenstein's Forms of Life, Patterns of Life and Ways of Living in *Nordic Wittgenstein Review*.

Narboux, P. (2004) 'Jeux de langage et jeux de dressage: Sur la critique éthologique d'Augustin dans les *Recherches philosophiques*' in *Ludwig Wittgenstein*, ed. C. Lecerf, *Europe*, 130-42.

Nelson, K. (2009) 'Wittgenstein and contemporary theories of word learning' in *New Ideas in Psychology* 27 (2009) 275–287.

Newport, E.L., Gleitman, H., & Gleitman, L.A. (1977) 'Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style'. In Snow & Ferguson (Eds.) *Talking to Children: Language input and acquisition*. Cambridge: Cambridge University Press, 109-49.

Ramscar, M. & Yarlett, D. (2007) 'Linguistic self-correction in the absence of feedback: A new approach to the logical problem of language acquisition' *Cognitive Science*: 31, 927-60.

Ramscar, M., Dye, M., & McCauley, S. M. (2013) 'Expectation and error distribution in language learning: The curious absence of "mouses" in adult speech' *Language* 89:4, 760-93.

Rorty, R. (2004) *The Brain as Hardware, Culture as Software*. *Inquiry* 47, 219-35.

Saxe, R. (2008) *The Ape That Teaches*. Interview by PBS Nova. 1.1. 2008. Online.

Schoneberger, Ted (2010) 'Three Myths from the Language Acquisition Literature' *The Analysis of Verbal Behavior*, 26: 107–131

Singleton, D & Ryan L. (2004) *Language Acquisition: The Age Factor*, 2nd ed. (Clevedon, UK: Multilingual Matters).

Tomasello, M.

---- 1999 *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.

---- 2008 *Origins of Human Communication*. Cambridge, Ma: MIT Press.

Yarlett, D & Ramscar, M. (2008) *Language Learning Through Similarity-Based Generalization*. Unpublished manuscript.

Young, M. W. (ed.) (1979) *The Ethnography of Malinowski: The Trobriand Islands 1915-18* (London: Routledge & Kegan Paul).

Wittgenstein, L.

BB *The Blue and Brown Books*, 2nd ed. Oxford: Blackwell, 1969.

BT *The Big Typescript: TS 213*, ed. & trans. by C. Grant Luckhardt & Maximilian A.E. Aue. Blackwell Publishing, 2005.

CE 'Cause and Effect: Intuitive Awareness' in *Philosophical Occasions: 1912-1951* ed. J.C. Klagge and A. Nordman. Indianapolis: Hackett Publishing, 1993, 371-426.

LPP *Wittgenstein's Lectures on Philosophical Psychology 1946-47*, notes by P.T. Geach, K.J. Shah and A.C. Jackson, ed. P.T. Geach. Hassocks: Harvester Press, 1988.

MWL Moore's *Wittgenstein Lectures in 1930-1933* in *Philosophical Occasions: 1912-1951*, ed. J.C. Klagge and A. Nordman. (Indianapolis: Hackett Publishing, 1993) 46-114.

- NL 'Notes on Logic', Appendix 1, in *Notebooks 1914-16*, ed. G.E.M. Anscombe and G.H. von Wright, tr. G.E.M. Anscombe, 2nd ed. Oxford: Blackwell, 1979, 93-107.
- OC *On Certainty*, eds G.E.M. Anscombe and G.H. von Wright, tr. D. Paul and G.E.M. Anscombe, amended 1st Edition. Oxford: Blackwell, 1977.
- PG *Philosophical Grammar*, ed. R. Rhees, tr. A. Kenny. Oxford: Blackwell, 1974.
- PI *Philosophical Investigations*, tr. G.E.M. Anscombe, 2nd Edition. Oxford: Blackwell, 1997
- PLP *The Principles of Linguistic Philosophy* by F. Waismann, ed. R. Harré, 2nd Edition. London: Macmillan, 1997.
- PR *Philosophical Remarks*, ed. R. Rhees, tr. R. Hargreaves and R. White. Oxford, Blackwell, 1975.
- RFM *Remarks on the Foundations of Mathematics*, ed. G.H. von Wright, R. Rhees and G.E.M. Anscombe, tr. G.E.M. Anscombe, 3rd revised Edition, Oxford: Blackwell, 1978.
- RPP I *Remarks on the Philosophy of Psychology*, vol. I, ed. G.E.M. Anscombe and G.H. von Wright, tr. G.E.M. Anscombe, Oxford: Blackwell, 1980.
- RPP II *Remarks on the Philosophy Of Psychology*, vol. II, ed. G.H. von Wright and H. Nyman, tr. C.G. Luckhardt and M.A.E. Aue, Oxford: Blackwell, 1980.
- TLP *Tractatus Logico-Philosophicus*, tr. D. F. Pears and B.F. McGuinness, London: Routledge & Kegan Paul, 1961.
- Z *Zettel*. G.E.M. Anscombe and G.H. von Wright (eds). Trans G.E.M. Anscombe, Berkeley: University of California Press, 1970.