*Twangling Instruments: Is parametric variation definitional of human language?**

NEIL SMITH & ANN LAW

Be not afeard: the isle is full of noises,
Sounds and sweet airs that give delight and hurt not.
Sometimes a thousand twangling instruments
Will hum about mine ears; and sometimes voices …
[Caliban in *The Tempest* Act III Scene 2]

Abstract

We investigate the hypothesis that what is unique to human language is not recursion but the existence of parametric variation (PV), where this is a way of formulating and uniting two problems: Plato's problem and the limits of typological variation. We suggest identity criteria for parametric variation in language, and see if any of the properties generalise first to other human cognitive domains - particularly music and morality, and second to animal cognition, especially birdsong. We consider several possible outcomes:

a. PV is unique to human language
b. PV is unique to humans but not just to language
c. PV is common to human language and birdsong, but not the rest of cognition
d. PV is common to everything – language, cognition, birdsong …
e. There is no coherent (or uniform) notion of PV.

Our tentative conclusion is (a) above.

1 Introduction

What makes human language unique? For Hockett (1958) and Hjelmslev (1961) it is 'double articulation'; for Jackendoff et al. (2006) it is a rich vocabulary; but the most influential recent suggestion is recursion: either recursion *tout court* or 'phase recursion' (Postma & Rooryck 2007). Recursion has featured prominently since Hauser et al (2002) proposed as an empirical hypothesis that what is unique to

* This paper is a revised version of one presented in Utrecht in April 2007 at the workshop on "Birdsong, Speech and Language: Converging Mechanisms?" For comments, conversation and correspondence we are grateful to Misi Brody, Noam Chomsky, Annabel Cormack, Tecumseh Fitch, Paul Harris, Gary Marcus, Rita Manzini, Amahl Smith and Deirdre Wilson. None of these should be taken as agreeing with anything we say. As we are still in the process of making further revisions any reaction to this instar would be welcome.

language and unique to humans is '(probably) recursion'.[1] This answer is (probably) wrong: recursion is not unique to language, but is characteristic of the Language of Thought (in Fodor's 1975 sense; cf. also Fitch et al, 2005, Smith, 2004) and it may not be unique to humans given the hierarchical structure of canary song (Gardner et al, 2005), the improvisation found in whales (Payne, 2000:135), and perhaps the properties of animal navigation. Another potential answer that we investigate is parametric variation (PV) where this is a way of formulating and uniting two linguistic problems: the puzzle of first language acquisition (Plato's problem, Chomsky, 1986) and the limits of typological variation. Despite Hauser's (2006) spirited promotion of the claim that PV is characteristic of moral judgement, and Smith's (2007) parallel suggestion for music, we think it is plausible to suggest that PV is unique to human language and that the variation found in other cognitive domains and in animal vocalisations is not 'parametric'. We do not wish to exclude other possibilities: the putative uniqueness of PV to human language may be derivative from other characteristics, such as an immensely rich lexicon, and the uniqueness of human language itself may well reside in the constellation of a number of different properties.

The structure of the rest of the paper is as follows: in section 2 we outline the properties of PV in language, concentrating on the difference between parametric and non-parametric variation; in section 3 we see if these properties generalise to other domains of human cognition, in particular morality and music; and in section 4 we then see if they generalise to domains of animal cognition, in particular birdsong. Finally, in section 5 we entertain a number of alternative conclusions as in (1):

(1) a. PV is unique to human language
b. PV is unique to humans but not just to language
c. PV is common to human language and birdsong, but not the rest of cognition
d. PV is common to everything – language, cognition, birdsong …
e. There is no coherent (or uniform) notion of PV.

For historical reasons we exclude logical possibilities such as (f):

f. PV characterises e.g. birdsong but **not** human language

We tentatively endorse (1a) and suspect that, if PV is definitional of human language, it is because PV is a solution to Plato's problem and only human language confronts the learner with this problem in its full complexity. Irrespective

[1] It is not easy to interpret such claims precisely: Tomalin, 2007, points out that 'recursion' is used in at least five different ways so that evaluating Hauser et al's (2002) claim is problematic.

of the validity of restricting PV to human language it is clear that each of the domains discussed is regulated by its own universal principles.

2 Parametric variation in the language domain

2.1 'Principles and Parameters' theory

PV is part of 'Principles and Parameters' theory (Chomsky, 1981a; for overviews and history see: Williams, 1987; Roberts, 1997; Baker, 2003). The human language faculty is standardly described in terms of a contrast between the Faculty of Language in the 'broad' sense (FLB) and a proper subpart of that faculty referred to as the Faculty of Language in the 'narrow' sense (FLN) (Hauser et al, 2002). The former includes a variety of performance mechanisms for parsing and producing utterances as well as our strictly grammatical ability. Many parts of FLB are shared with other organisms from bumblebees to bonobos, but FLN – which may be empty - is by hypothesis unique to humans and unique to language. FLN is characterised in terms of Universal Grammar (UG), the innate endowment which allows children to learn their first language and which defines the basic format of human language. It specifies that human languages consist of a Lexicon and a 'Computational system' (referred to as C_{HL} – the computation for human language). The lexicon consists of a set of lexical entries, each of which is a triple of phonological, morpho-syntactic and semantic features, and with a link to associated encyclopaedic information. Thus a word like *bumblebee* is specified as being stressed on the first syllable (phonological information), consisting of, two elements *bumble* and *bee* (morphological information), being a (count) noun (syntactic information), referring to an insect of the genus *Bombus* (semantic information). The link to encyclopaedic memory will relate this entry to such idiosyncratic information as that bumblebees buzz, are large and hairy, etc. Every natural language has a lexicon containing tens of thousands of such entries whose essential function is to link representations of sound to representations of meaning.

UG also provides a set of exceptionless principles such as structure dependence (Chomsky, 1971), (strict) cyclicity (Freidin, 1999; Chomsky, 2002), the Extended Projection Principle (Chomsky, 1995), etc. which constrain the operation of the computations. Structure dependence is a principle which states that all grammatical operations – phonological, morphological and syntactic – have to be defined in all languages over structures rather than simple linear sequences of element. That is, the possibility of counting the number of words or constituents is excluded *a priori*. For instance, the formation of a question from a congeneric statement as in (2) can refer to syntactic categories, such as 'auxiliary verb', and their movement to a particular position (the Infl node), but not to the 'third word'. In (2) the effect might appear to be the same, but the more complex (3), with the two possible questions in (4), shows that only one process is licit.

(2) The man is in the room – Is the man in the room?

(3) The man who is tall is in the room

(4) a. Is the man who is tall in the room?
b. *Is the man who tall is in the room?

It is significant that this principle acts as a constraint on language acquisition: children learning their first language have their 'hypothesis space' constrained with the result that they never make mistakes like that in (4b).

The extended projection principle (EPP) stipulates that all clauses must have a subject, so (5a) is acceptable but (5b) is impossible:

(5) a. John came home early
b. *Came home early

The status of this principle is somewhat different from that of structure dependence in that some 'pro-drop' languages seem to be systematic exceptions. As we shall see, this exception is only apparent.

In addition to a set of universal principles, UG also provides a set of parameters which jointly define the limits of variation. This is typically conceptualised as the setting of a number of 'switches' – on or off for particular linguistic properties. Typical examples of such parameters in syntax are the head-direction parameter (whether heads, such as Verb, Noun and Preposition, precede or follow their complement), the null-subject (or 'pro-drop') parameter (whether finite clauses can have empty pronominal subjects), and the null-determiner parameter (whether noun phrases can have empty determiners). English and Hindi have opposite values for each of these parametric choices, as illustrated in (6) to (8):

(6) Head-first – "on the table" – Head-last "mez par"
table on

(7) Non-pro-drop – *"Is working" – Pro-drop – "Ø kaam kartaa hai"
work doing is

(8) Non-null D – *"boy has come" – Null D – "Ø laṛkaa aaya hai"
boy come is

Typical examples in phonology are provided by the stress differences characteristic of English and French, and the possibility of complex consonant clusters found in English but not in Japanese. English stress is 'Quantity-sensitive' whereas French

stress is 'Quantity-insensitive', with the result that words with the same number of syllables may have different stress in English but uniform stress in French, as shown in (9):

(9) *América / Manitóba // endurcissement / sentimental*

In English words may begin with clusters of consonants in a way which is impossible in Japanese, with the result that English loans into Japanese appear with the clusters separated by epenthetic vowels, as shown in (10):

(10) *screwdriver // sukuryūdoraibā*

The theory unifies two different domains: typology and acquisition. Variation among the world's languages (more accurately the set of internalised I-languages, Chomsky, 1986) is defined in terms of parametric differences such as whether verbs precede their objects as in English, or vice versa as in Hindi. In first language acquisition the child's task is reduced to setting the values of such parameters on the basis of the stimuli it is exposed to – utterances in the ambient language. Given the strikingly uniform success of first language acquisition, "the set of possibilities [must] be narrow in range and easily attained by the first language learner" (Smith, 2004:83). By hypothesis, the principles do not vary from child to child or from language to language: as Chomsky (2006:183) puts it: "acquisition is a matter of parameter setting, and is therefore divorced entirely from … the principles of UG".

The theory is at once 'internalist' (i.e. it is a theory of states of the mind/brain) pertaining to knowledge which is largely unconscious, and universalist. An immediate implication of this position is that the range of parametric choices is known in advance and, as a corollary, it claims that acquisition is largely a process of 'selection' rather than instruction and that such acquisition is likely to take place in a critical period or periods.

The claim that PV is unique to human language could be trivially true – PV presupposes UG and – by hypothesis – neither animals nor other human domains have UG. However, Hauser (2006) and Mikhail (2007) have both proposed a 'Universal Moral Grammar' akin to UG[2], and Hauser is explicit in his defence of PV in the moral domain. Similarly, Trehub & Hannon's (2006:82) observation that infants react to and learn music spontaneously and rapidly is compatible with a parameter-setting account. In any case, we attempt to make the claim non-trivial by extracting from the wealth of views in the literature some putative identity criteria for PV and seeing if they generalise. That such an extension is plausible is

[2] For instance Hauser (2006:419-420) writes: "Underlying the extensive cross-cultural variation we observe ... is a universal moral grammar that enables each child to grow a narrow range of possible moral systems".

suggested by a number of intuitive commonalities between language, music and morality on the one hand, and between language and birdsong on the other. In all areas there are putative universals, suggesting an innate basis; there are simultaneously clear effects of the environment, suggesting an interplay between genes and learning; and there are parallels in the various stages that organisms pass through in mastering the complexity of the system they are acquiring. We shall claim nonetheless that PV is indeed unique to human language. We look first at PV in syntax and then cast a brief look at PV in phonology.

2.2 PV in Syntax

The main discussion of PV has been in the domain of syntax. Originally, parameters were associated with the principles of UG but following Borer's (1984) work were later located in the lexicon (see Smith, 2004, for discussion). As Kayne puts it: "syntactic parameters are … necessarily features, or properties, of elements of the lexicon" (Kayne, 2005:4). Moreover, for Chomsky and many others (e.g. Chomsky, 1995), the relevant part of the lexicon is restricted to the functional lexicon, where this refers to that subset of the whole which deals with functional categories such as Tense, Complementisers and Determiners, in contradistinction to the conceptual lexicon which deals with substantive categories such as Noun, Verb, Adjective and (perhaps) Prepositions. Thus *bumblebee,* illustrated above, belongs to the conceptual lexicon, whereas an item such as *the* belongs to the functional lexicon. Two major differences are associated with this distinction: members of the functional lexicon characteristically have no encyclopaedic link (except, for the literate, their spelling) and, more importantly, may also lack any phonological content. As illustrated above, some languages (like Latin or Russian) allow empty determiners, others (like Spanish or Greek) allow empty (null) subjects, whereas others (like English and German) allow neither of these possibilities. Such differences are a matter of PV. In contrast, a cross-linguistic difference in the subcategorisation properties of a verb such as *convince*, where *John convinced Mary to go* is grammatical (for some) or ungrammatical (for us) is a matter of (non-parametric) variation in the conceptual lexicon.

In the 'old' system the pro-drop parameter was a possible setting of the EPP; in the 'new' system the pro-drop parameter is instantiated as the condition that (the functional category) AGR = PRO (for pro-drop languages) *versus* AGR ≠ PRO (for non-pro-drop languages). Recently, with the demise of 'AGR' in Minimalism the form of the parameter might be: "[Spec, IP] is obligatory". In either case, pro-drop is a property of particular functional categories.

There is a general consensus on the need for such a (traditional) distinction between functional and substantive categories, but little agreement on how to draw the boundary lines between the two (for discussion, see Muysken, in press). The account of functional categories which is closest in spirit to ours is that provided by

Roberts & Roussou (R&R) (2003:28) who define them on the basis of their behaviour at each of the two levels of representation common to all theories: LF and PF. LF represents the interface with the conceptual-intensional system – i.e. where the language faculty connects with the representation of thought, PF represents the interface with the sensori-motor system where the language faculty connects with systems of audition and articulation. For a linguistic representation to be well-formed all its constituent elements must be interpretable at one or both of these interfaces. Lexical items are designated +p and +l to indicate that they must have an interpretation at PF or LF respectively. R&R define functional categories as "that class of syntactic categories which is not obligatorily +p", whereas lexical categories such as Noun and Verb are always +p, +l. In other words lexical categories must be interpretable at both interfaces, but functional categories such as C(omplementiser), T(ense), D(eterminer) may be overt or covert. If they are 'covert' or empty, they have no interpretation at PF, they are syntactically and semantically present but are not pronounced.[3] The set of interpretable features is provided as a set of substantive universals by UG. A functional head, such as D, is then marked for some feature and will be further specified in a language-particular fashion as having this feature manifest overtly or covertly. The relevant functional category will then be realised in the syntax by either Merge or Move (R&R, 2003:30). To make the discussion concrete they illustrate with a simple example of differences among yes/no questions, marked syntactically by the presence of an abstract morpheme Q. In colloquial French Q is unmarked and questions are indicated just by intonation (*Jean a vu Marie?*). In Welsh, Q is marked and also overt so a question particle *a* is merged in initial position (*A welodd John Mary?*). In English Q is marked and covert so the auxiliary moves to the front (*Did John see Mary?*). Many theories restrict PV to a choice between the values [+/- strong] of functional categories. For such theories parameters are necessarily binary, but such a restriction is independent of the general conceptual claims of the theory.

A further difference between competing theories of PV revolves around the deductive consequences of the choice of a particular value for some parameter. This difference is frequently labelled as a distinction between 'macro-parametric' and 'micro-parametric' variation. 'Macro'-PV is typically exemplified by the head-direction (head first/head last) parameter (Chomsky, 1981a) or Baker's (1996) polysynthesis parameter which determines the overall morphological structure of the language. Each of these parameters has a wide variety of effects, whereas 'micro-PV' – of the sort exemplified by the choice of auxiliary to accompany unaccusative verbs (Perlmutter, 1978; Burzio, 1986) or case realignment in Albanian causatives (Manzini & Savoia, 2007) is characteristically more restricted and has correspondingly fewer repercussions. Roberts (1997:273) lists a number of

[3] Though they may have phonological effects. See Smith, 2005, ch. 15, for discussion.

such parameters and their interrelations, though without adverting to the macro-/micro- distinction.

For our purposes, the distinction is not crucial, though it *is* important for us to determine what constitutes PV and what is non-parametric, a distinction that is again a matter of controversy in the literature. Many, perhaps most researchers deny a principled distinction between macro-, micro- and non-parametric variation. For instance, Kayne is sceptical of validating a distinction between micro- and macro-parametric variation. Indeed, he suggests that "[e]very parameter is a microparameter" (Kayne, 2005:10), and probably "[e]very functional element made available by UG is associated with some syntactic parameter" (p.11) – perhaps never more than one (pp.14-15). He suggests about 100 functional elements. Further, if a language "visibly has some functional element, then all languages must have it" (p.16). The problem is to decide what the features involved which must be "simple and limited in type" (ibid) can be. Kayne uses the term to "characterize all cross-linguistic syntactic differences" (ibid p.6) and exemplifies the concept with the contrast in word order between the English and French pairs: *too rich/rich enough* versus *trop riche/assez riche*. Similarly, Manzini & Savoia (2007, Rita Manzini – pc) deny any distinction both between macro- and micro-parametric variation and also between parametric and non-parametric variation. Their respective positions differ in that while Kayne arrogates all change to the functional lexicon by postulating large numbers of abstract (silent) categories (cf. Roberts & Roussou, 2003:24), Manzini & Savoia do it by treating all variation on a par, as simply 'lexical', with no special status for any functional domain, no silent categories and, crucially, no non-parametric variation.

We have two reactions to this position: first, we believe that it is too broad, as not all syntactic differences seem to us to merit the description 'parametric'; second, we are not convinced that the contrast between macro and micro should be given up so easily. Baker argues that there is no principled difference among types of PV – micro-PV is that which is "localized to a particular configuration" whereas macro-PV shapes "every phrase of every clause", but we share Chomsky's intuition that "there seems to be a major distinction between macroparameters like the head parameter … and what some call 'microparameters'", and his "doubt that there will ever be much of a theory about these micro-variations" as they are "basically accidents" – "what a person cannot know in advance" (pc December 06). We are happy to remain non-committal on the 'macro-/micro-' dimension, but it is important that we differentiate the parametric and the non-parametric and we will attempt to give identity criteria for PV.

2.3 PV in Phonology

The domain of PV in syntax is, crucially, the set of functional categories, but there is no comparable constraint in phonology. Rather, as indicated in the examples of

stress and possible onset clusters above, parameters are associated with words, syllables, vowels, etc. (cf. Dresher & Kaye, 1990, Dresher, 1999). It follows that PV is not restricted to the functional lexicon and could be operative in areas without such a category. Accordingly we need to identify more abstract properties of the concept. A crucial property of PV is that it gives rise to a situation in which language acquisition is cued or triggered rather than 'learned'. Learning in the traditional psychological sense (i.e. a process involving hypothesis formation and testing, association, generalisation, and so on) plays little role in first language acquisition. As Chomsky (1980:134) put it: "in certain fundamental respects we do not really learn language; rather, grammar grows in the mind." Language acquisition, however, is clearly dependent on external input, and this input is said to trigger or cue linguistic development. We think that the notion 'cue' is likely to be more fruitful than 'trigger', but the basic distinction is that between either of these and traditional learning.

2.4 Typology and acquisition

We need to recapitulate a little. PV is variation within a narrow range defined by universal principles – languages are not like Joos (1957:96) suggested: they cannot "differ from each other without limit and in unpredictable ways". It distinguishes among possible mentally represented states of the language faculty: i.e. it defines the range of possible I-languages (Chomsky, 1986) and makes available a typology of the world's E-languages. PV unites – as an empirical claim - two domains: typology and first language acquisition (cf. e.g. Fanselow, 1993:xvi), but it is important to note that acquisition has conceptual priority, with the typological exploitation of the claims being derivative. If it is correct to claim that first language acquisition consists in setting parametric choices these must have priority over any taxonomy based upon them. The typological claims being made are nonetheless not trivial, accounting as they do for the interdependence of particular properties across languages: e.g. the fact that SOV, PostP, Suffixation, etc. cluster together. If acquisition is a matter of parameter setting and if there is no negative evidence available to the child then all the possible alternatives are antecedently known or innate[4] and the child's task in learning its first language is a matter of selecting a grammar on the basis of the particular properties of the input, rather than needing instruction (cf. Piattelli-Palmarini, 1989).

This claim of 'antecedent knowledge' or 'knowledge without experience' has a number of implications. The first of these is that parametric choices may give rise

[4] The term 'innate' is problematic because of different usages (see Marler, 1999, for discussion). An obvious contrast between (human) language and e.g. birdsong is that there are no known linguistically relevant genetic differences between humans, whereas there are known and experimentally manipulable differences among different populations of canaries (Mundinger, 1999).

to cascade effects: coming to know one fact (e.g. that Verbs precede their objects) licenses knowledge by the learner of other facts (e.g. that Prepositions precede their objects) without further exposure. 'Cascades' have become unfashionable because of the dissociation of the properties associated with (notably) the pro-drop parameter (for discussion, see Ackema et al, 2006). This reaction may have been hasty: cascades could be operative in the domain of acquisition even if there is such dissociation. That is, the child leaps to the 'cascade conclusion' (i.e. selects one parameter on the basis of the setting of a distinct but related parameter) unless there is evidence to the contrary, thereby solving Plato's problem (in part) (for discussion, see Smith & Cormack, 2002).[5] A second implication of 'selection' is that parametric choices allow no analogical formations ("analogy works only for non-parametric choices", Smith, 2004: 121), and license only a subset of the logically possible mistakes a learner could make. Further, as pointed out originally by Hyams (1987:18), this conception also makes available a wider selection of triggering data, and gives rise to the kind of 'network' described by Roberts (1997:275) where parameters are not independent of each other. The details are complex, contentious and of only limited relevance to our current concerns, so we turn next to a more detailed discussion of the distinction between parametric and non-parametric variation.

2.5 Identity criteria for parametric variation

The theory of PV stipulates that the range of choices is 'antecedently known', and this basic property, our first criterion, correlates with a number of others which distinguish PV from non-parametric variation, and allow us to provide identity criteria for PV. Our second criterion is that variants licensed by parametric choice must be cognitively represented. Consider acclimatisation, specifically sweating. We have a critical period for setting our sweating switch: experiencing hot and humid weather in the first three years of life leads to a different setting from exposure to different conditions. These settings cannot be significantly altered after the critical period. Despite a certain superficial similarity, this is not PV because the different states are not (mentally) represented and have no cognitive effects. Comparable remarks obtain with regard to Hauser's (1997) discussion of the notion 'deception'. It is perhaps mildly surprising to learn that rhesus monkeys practise systematic deception but the evidence is cogent. It is less plausible to think that stomatopods (shrimps) practise it simply because we do not attribute to them the 'cognitive processes' necessary for deceit, even if their behaviour could be interpreted as deceptive if they were intellectually comparable to us in other respects. For the same reason, Hauser's 'sceptical honeybees' are probably best

[5] A clearer example of a cascade may be provided by the contrast 'Signed/Spoken' where choice of the value 'Signed' entails a range of other choices such as +Pro-Drop. In fact, the variation in the syntax of signed languages seems to be much narrower than in spoken languages.

treated behaviourally rather than cognitively. Differences in neurobiology and the absence of relevant experimental results make it difficult to determine where songbirds come in this hierarchy.

Our third criterion is systematicity. A simple example is provided by irregular morphology of the type exemplified by the impossibility of **amn't* in (most varieties of) English, or the kind of defective paradigm seen in Latin *vis-vim-vi.* We do not consider this to be PV because it is by definition not systematic and hence we could not plausibly acquire knowledge of it by any process of triggering. Our fourth criterion is dependence on the input. An example is provided by the individual variation in e.g. consonant harmony in phonological development (cf. Smith, 1973:163), or the variation in the choice of initial or final negation in syntactic development (cf. Smith, 2005:29). For instance, two children in essentially the same environment may produce the adult *duck* as [gʌk] and [dʌt] respectively. These are both manifestations of consonant harmony, but they do not count as PV because the particular variants chosen appear to be independent of the input. There is, of course, some dependence on the input: the form of the vowel, the absence of a labial articulation in either initial or final position, but the choice of [gʌk] or [dʌt] is idiosyncratic to a particular child. It instantiates a general strategy (consonant harmony) which consistently produces variants none of which occurs in the adult language. 'Dependence on the input' thus stipulates that parametric choices must reflect possibilities in the target (ambient) language. To give a syntactic rather than a phonological example, the development of negation typically goes through a stage in which the negator is peripheral, either initial or final. Individual children then differ such that one child learning English may say '*no like cabbage*' and another '*like cabbage no*'. Again we take this variation to be non-parametric, even though it may be hard to differentiate it from UG-licensed errors of the sort described by Crain and his colleagues (cf. Crain & Pietroski, 2002), where a child produces a form which never occurs in the input (e.g. '*What do you think what pigs eat?*') because the structure is licensed by UG and so occurs in other languages.[6] To our knowledge, no language allows peripheral negation with the negator either initial or final. Despite this potential difficulty, we think that the case of consonant harmony in phonology and negation in syntax makes the conceptual contrast between parametric and non-parametric variation clear.

Our fifth and final criterion is that PV must be deterministic: that is, the input to the child must be rich enough and explicit enough to guarantee that a parameter can be set. If the input does not meet this requirement we are dealing with non-parametric, random, variation. This criterion arises from a reinterpretation of Smith & Cormack's (2002) discussion of 'parametric poverty'. They suggested that there are 'random settings': given the same input, different children might assign the

[6] We take it that such over-generalisation is a sign that the child has, temporarily, mis-set the relevant parameter.

same parameter different values. They illustrate this with the possibilities for 'sequence of tense' phenomena. For many speakers *Did you know that Emily is ill?* is simply ungrammatical, and only *Did you know that Emily was ill*" is possible; for us both are fine, though with interpretational differences. Smith & Cormack suggest that the parameter has been fixed at random; we prefer to think that such 'randomness' reveals that the variation is non-parametric. Comparable examples can be found in phonology: Moira Yip (2003:804) gives an example where some people treat a post-consonantal glide as a secondary articulation of the consonant, others as a segment in its own right: "the rightful home of /y/ [is] underdetermined by the usual data, leaving room for variation". Her conclusion is that "speakers opt for different structures in the absence of conclusive evidence for either". Again that indicates for us that the variation is non-parametric.

In Table 1 we summarise and exemplify these criteria, all of which are common to syntax and phonology.

1. The range of choices must be antecedently known; hence acquisition is a matter of selection rather than instruction	
Parametric Movement; ellipsis	Non-parametric Irregular morphology; lexical exceptions (*likely/probable)*
2. Parametric choices must be mentally represented	
Parametric Stress; word-order.	Non-parametric Sweating; consonant harmony (e.g. 'duck' ⇨ [gʌk] or [dʌt])
3. Choices must be systematic - variations are not accidents	
Parametric Wh-movement (covert or overt)	Non-parametric Defective paradigms (**amn't*; Latin *vis-vim-vi*)
4. Choices must be dependent on the input and hence correspond to a possible state of the adult language	
Parametric Quantity-sensitivity Word order – head direction	Non-parametric Consonant harmony Early negation (*no computer on/ computer on no*)
5. Choices must be deterministic	
Parametric Pro-drop Complex onsets in phonology	Non-parametric Sequence of tense Post-consonantal glides

Table 1: Identity criteria for Parametric Variation

The restriction to the functional lexicon of syntactic variation[7] is an epiphenomenon of the fact that PV operates over elements of the appropriate domain: the units manipulated in the syntax are words (i.e. elements of the lexicon which appear in the numeration); the units manipulated in the phonology are syllables and feet (and sometimes phonological words). The criteria are intended to be jointly sufficient and individually necessary to identify PV. We should add that some phenomena (e.g. cascade effects) may be sufficient to license the conclusion that there is PV even though it is not possible to make this a necessary condition. Hence we need to distinguish 'cascade effects' - knowledge without *further* input - from the fixing of a form as in consonant harmony[8] or peripheral negation where the input does not provide any basis for the choice. Assuming the continuity hypothesis (Pinker, 1984) and the idealisation to "instantaneity" (Chomsky, 1981b:224) this fixing reduces to the claim that the grammar the child has attained must generate forms of the ambient language. As a final point it should be noted that occurring in a critical period is compatible with PV, but is not evidence for it.

Assuming that, despite the different interpretations available in the literature, the nature of PV is clear, we now proceed to see if and how it generalises. We shall look in each case at the range of variation involved to see whether it is comparable to parametric variation in language, and whether the acquisition of the relevant ability is reminiscent of language acquisition.

3 Generalisation of Parametric Variation in the domain of human cognition

It might appear most plausible to generalise PV to areas which, like language, rely on auditory input: music and the analysis of environmental sounds (noise). However, the parallels may not be quite as obvious as this suggests: music and noise are necessarily tied to the auditory channel whereas signed languages show that the language faculty is not restricted in this way. We look briefly at the analysis of noise, before concentrating on moral judgement and music.

The case of 'noise' is interestingly parallel to language. Smith (1999) observes that dissociation patterns in the various agnosias (cortical deafness, auditory agnosia, pure-word deafness and phonagnosia) reveal remarkable richness of structure, and speculates that growing up in different auditory environments (a jungle, by the sea, in the desert) might result in the setting of different parametric values for auditory discrimination. Work by Saygin et al (2005) shows an appreciable amount of parallelism between the two domains but also some notable differences, (e.g. in the brain areas involved) and the authors are sceptical that PV

[7] It is possible that some variation – e.g. that for which the subset principle and markedness are irrelevant – is still non-lexical. (cf. Safir, 1987:87).

[8] If Smith, 2005 ('gucks'), is correct, consonant harmony would not be mentally represented.

is at work (Fred Dick, p.c.). In the absence of solid ontogenetic evidence we leave the subject for future work.

3.1 Moral judgement

Moral judgement is not an axiomatic system with the deductive structure of a theory of language and is correspondingly resistant to formalisation. However, there are striking parallels, including such properties as being 'internalist', exploiting a competence/ performance distinction, consisting in knowledge which is largely unconscious, being acquired under poverty of the stimulus[9] - triggered rather than learned; universal – hence uniform and innate; and dependent in part on the development of a Theory of Mind (ToM) module (cf. Papafragou, 1998). Building on Chomsky's insights about language and his suggestions about morality (cf. Chomsky, 1975), e.g. that "Hume ... pointed out that the foundation of morals must be what we nowadays call generative grammar" (Chomsky, 2003:40), Hauser (2006, cf. Hauser et al, 2007) and Mikhail (2007) have elaborated an analysis of moral judgement which has many of the hallmarks of our knowledge of language. After postulating a universal principle of justice as 'fairness' Hauser states that "[a]dopting the analogy to language, one would expect a universally held principle of fairness that varies cross-culturally as a function of parametric variation; experience with the native environment triggers the culture's specific signature of fairness and fair exchange" (2006:72). More worryingly, he writes that "[i]n the same way that our universal grammar provides a toolkit for building a specific grammar, in which certain principles and parameters hold and others do not, our universal moral grammar provides a different toolkit, enabling us to implement particular principles and parameters but not others" (2006:74). This formulation is unfortunate in that it suggests that the *principles* can vary from culture to culture rather than just the parametric choices.

Hauser's statement is inconsistent in claiming that Universal Moral Grammar works "in the same way" as UG, but allowing that certain principles can vary. Interestingly, he also raises the question whether parametric moral cascades occur, asking "Do moral systems work like language in the sense that choosing to set certain parameters influences subsequent settings?" (2006:300). He is also explicit that his position is one where the moral system children construct "depends upon their local culture and how its sets the parameters that are part of the moral faculty" (2006:303).

Apart from Hauser's heterodox interpretation of principles and parameters, there are also systematic asymmetries between language and morality which lead to doubts about the validity of the parallel between the two domains. First, the building blocks of moral judgement are actions, causes, consequences and their

[9] This may be contentious, but the stimuli which give rise to our largely unconscious moral judgments may be concealed in the linguistic input; see Smith, 2007, for discussion.

arrangement over a period of time. Second, the proprietary lexicon of the moral system is not comparable to that of a natural language. That is, one might expect that any cognitive module should have its own lexicon disjoint from that of any other. However, once one goes beyond universal principles such as 'justice as fairness' (see below) the units of moral judgement are a subset of the items in the conceptual lexicon of the culture the system is embedded in. Such systems dispose of a huge set of concepts from the language of thought and these are subject to variation, but the variation lies more in the detail of encyclopaedic knowledge and is idiosyncratic rather than parametric. For these reasons we are sceptical of the force of Hauser's suggestions: we think the typological claims may be plausible but that there is no evidence for PV in the domain of acquisition. Although there is some literature on children's development of moral judgement (e.g. Harris, 2000) there is almost no systematic cross-cultural work on anything that might be considered parameter-setting. Accordingly, we look first at some putative universal principles, as in (11), and specific cultural parameters, as in (12), and evaluate how close they are to what is found in language. In a nutshell, we are happy with Hauser's claims about the existence of universal principles (though these should not be subject to cross-cultural variation) and their parallels with what is found in language; we are not convinced that he has made a case for PV in anything like the current linguistic sense.

(11) Universal moral principles:

a. Justice as 'fairness' in the distribution of advantages. Hauser (2006:83) states that "fairness is a universal principle with the potential for parametric variation and constraints". Following Lakoff (1996) Hauser (2006:75) suggests that such variation might be exemplified by the difference between equal or scalar distribution of responsibility ('effort is equitably shared' *versus* 'those who can do more have greater responsibilities') or a needs-based *versus* scalar distribution of rewards ('those who need more, get more' *versus* 'those who work harder, get more').
b. Obedience to authority
c. Incest taboo; nature of 'shame'.

(12) Putative cultural parameters of moral judgement:

a. There are clearly differences in the underlying philosophy motivating the fairness of distribution: for instance, a deontological versus utilitarian basis. This will be manifest in the existence of a social versus an individual basis of the kind (supposedly) characteristic of communism and capitalism. There will also be differences in the power

relations which are effective in any such distribution, as seen most clearly in the limits of the franchise. Everyone may have the vote, only men, only property-owners, only property-owning men over 21; only whites, only members of a particular religion, and so on.

b. Differences in the nature of authority: e.g. Theocratic/Secular, leading to further differences such as 'sin' versus 'crime'.

b´. Differences in the degree of obedience to authority, ranging from the anarchist to the totalitarian.

c. Definition of kin for incest; context in which 'shame' is felt.

The crucial issue for us is whether moral judgement satisfies the identity criteria for PV. Table 2 gives a summary overview.

1. Antecedently known	Dubious. The principles, such as 'Fairness' seem to be known in advance, but not the cultural variants.
2. Mentally represented	Yes
3. Systematic	In part: there is apparently systematic and non-systematic variation.
4. Dependent on the input	Yes but with differences between adult and child
5. Deterministic	Not consistently

Table 2 - Parametric Variation in moral judgement?

That the elements of the moral system are mentally represented seems straightforward. Whether this knowledge is antecedently known seems plausible for universal principles and the existence of categories such as 'good' and 'bad', but that parametric choices are available beforehand seems highly unlikely. Moral choices are usually systematic but as in language there is apparently arbitrary variation as well, as in the degree of 'moral disgust' felt by different individuals confronted with the same phenomena such as incest (see Jones, 2007). Here and more generally dependence on the cultural input is crucial. One's moral imperatives and one's judgements more generally are clearly likely to differ depending on whether one was schooled in a madrasa, a convent or a school in Beijing. The caveat expressed under 'dependent on the input' is prompted by the ontogenetic development of moral concepts, notably by the child's developing notion of lying. Children go through a stage in which they consider all untruths to be lies; only when they have developed a functioning Theory of Mind do they master the complexity of the adult concept, where some intent to deceive is necessarily included. That is, the child's use of the term 'lie' is unlike the adult's,

though it may correspond to something different in the adult language such as the phrase 'say something untrue'. This is unlike the kind of UG licensed error mentioned with regard to the Crain example above in the following respect. Children's moral misconception with regard to lying results from their representing only a subset of the information present in the adult form, whereas their syntactic error results from hypothesising a grammar partially disjoint from that of the adult language.

Two final points should be made: first, the unification of typology and acquisition seems not to be salient or even possible as yet. Second, the validity of the parallels Hauser draws in his (2006) relies on Chomsky's original notion of PV in which parameters were associated with principles of UG (*mutatis mutandis* with principles of UMG (Universal Moral Grammar), rather than the current one where they are associated with (features of) lexical items. In neither case would it be possible for the *principles* to differ from case to case in the way Hauser suggests. This makes comparison difficult, but on either interpretation PV seems to be different here than in language.

3.2 Music

We turn next to the scope for parametric analysis in the musical domain, looking at a number of properties characteristic of knowledge of music and its acquisition. First, we consider some obvious background differences between music and language. In music, the building blocks are (sequences of) notes arranged (as in language) into metrically articulated "hierarchical recursive structure" (Jackendoff & Lerdahl, 2006:38), there are fixed and discrete pitches for each mode, there is typically isochronicity, and there is no semantics. It is also striking that there is no musical equivalent of a natural language lexicon. It is true that Peretz & Coltheart (2003:690) refer to a 'musical lexicon', but this is characterised as "a representational system that contains all the representations of the specific musical phrases to which one has been exposed during one's lifetime" and a processing device designed to match incoming stimuli with stored representations. This makes it look like a parser and radically unlike the language lexicon. Given the point of departure of this paper, it is also worth putting on record that we do not find Jackendoff & Lerdahl's claim of recursion for music persuasive. What they describe is iteration rather than recursion.

Our focus of interest is the 'musical idiom' (Jackendoff & Lerdahl, 2006) and how hearers become 'familiar' with it. Specifically, "How does a listener acquire the musical grammar of MI [musical idiom – NS/AL]" and "What pre-existing resources in the human mind/brain make it possible for the acquisition of musical grammar to take place?" (ibid p.34). They also raise the issue of the cognitive specificity of the musical capacity which is explicitly parallel to the contrast between FLB and FLN, though their paper has surprisingly little discussion of any

of these issues, simply raising – and dropping – the observation that "each idiom will have its own characteristic structures, created out of the interaction of idiom-specific tonal and metrical principles with universal principles of tension and relaxation." (ibid p.58).

In looking at parallels and differences between language and music, the first observation to make is that the perceptual apparatus underpinning music is the same as that for language, and is moreover shared by some animals, and "the most parsimonious interpretation of the available evidence is that infant skills are a product of general perceptual mechanisms that are neither music- nor species-specific" (Trehub & Hannon, 2006:91). Only some skills, however: they point out that perception of relative duration is necessary to understanding rhythmic structure in both language and speech, even though these differ (ibid p.85), and musical meter is species-specific (ibid p.87). Moreover, as they emphasise (Trehub & Hannon 2006:82) infants react to and learn music spontaneously and rapidly in a way that animals do not.

This largely common perceptual foundation gives rise to a variety of properties shared by the two systems. Even though their status may be controversial (cf. Nettl, 2000, passim), there are clear musical universals – the use of octaves, scales of less than or equal to 7 pitches, the use of tonal scale systems, rhythm based on 2 or 3 (Brown et al, 2000:13-14; see also Peretz 2006). Moreover, as Nettl (2005:56-57) puts it: "People tend to absorb the fundamental grammars of their own language and their own music very early, to know very quickly whether a word belongs to their own language, and whether a particular interval or chord is proper in their own music". Similarly, "the compatibility of Western and traditional African music" (Nettl, 2005:58) suggests the operation of a typological parameter which is shared by these two areas. Nettl goes on (2005:67) to locate this compatibility in terms of harmony and draws a contrast in this respect between African and Middle Eastern musics.

In at least some cases (e.g. the octave) there may be a simple physical explanation for the universality, but there are interesting data on infants' sensitivity to different aspects of music (see e.g. Trehub, 2000) which are suggestive of parametric variation. Synchronised movement to music is universal (Trehub & Hannon, 2006:86); there are aspects of musical ability such as the possession of absolute pitch which seem to require exposure during a critical period (Brown et al, 2000:13); grouping, rhythm and meter are all underpinned by innate abilities (Trehub & Hannon, 2006:82), and musical grouping eventuates in a "hierarchical recursive structure" (Jackendoff & Lerdahl, 2006:38; but cf. the reservation above): that is, there is a putative 'cognitive homology' of metrical grids in language and music (Jackendoff & Lerdahl, 2006:42f.). Trehub & Hannon (2006:81) report that infants outperform adults on detecting contextually appropriate changes (a change in the correct key as opposed to an incorrect key – adults only detect the latter, infants detect both). This ability is strikingly parallel to the abilities of infants to

make categorial discriminations such as the l/r distinction which they lose unless the ability is reinforced by ambient input (cf. Jusczyk, 1997).

To make the discussion explicit, a specific example of what might constitute a parametric choice with obvious triggers for acquisition is provided by the variety of scales found around the world. There is a clear contrast among the Diatonic (Western heptatonic) scale, the Slendro (Indonesian pentatonic) scale used for the gamelan, and characteristic also of Chinese music, and the Śruti (Indian) scale with (usually) 22 śruti per octave.

The commonalities can be generalised. 'Musical idiom' is parallel to language not only in exploiting the same perceptual apparatus but in being 'internalist', largely unconscious, in becoming established in a critical period for (e.g.) absolute pitch, in showing the possession and loss of categorical discrimination (Trehub & Hannon, 2006; cf. Jusczyk, 1997), etc. But these similarities may be only superficial. Metrical structure is peculiar to music and language, whereas 'grouping' structure is also characteristic of other aspects of cognition.

Moreover, all this is neutral with regard to the question of PV in the acquisition of these abilities. However, Trehub & Hannon (2006) provide potentially relevant evidence for the generalisation of the typological to acquisition in their discussion of simple and complex meter: "After 6-month-old infants listen to a folk tune with a simple or complex meter for 2 min, they prefer variations that disrupt the original meter to those that preserve it both for simple and for complex meters ... 6-month-olds' differential responsiveness to meter-preserving and meter-disrupting variations parallels the ratings of Bulgarian and Macedonian adults. By 12 months of age, however, infants respond differentially to meter-preserving and meter-disrupting variations in simple-meter contexts but not in complex-meter contexts" (Trehub & Hannon, 2006:88). That this could be PV is suggested by the difference they note between this and other abilities: "the early acquisition of adultlike biases in perceiving rhythm and meter contrasts with the protracted developmental course of sensitivity to hierarchical pitch structure" (Trehub & Hannon, 2006:89).

We are now in a position to give in (13) and (14) a summary list of Universal principles and putative cultural parameters together with, in Table 3, an indication of whether these satisfy the identity criteria for PV.

(13) Universal musical principles:
   a. Tonal encoding of pitch
   b. Tonic as centre; hierarchical organisation of cognitive pitch space distances; tonal relaxation and tension
   c. Exploitation of Scales
   d. Exploitation of Metres

(14) Cultural parameters of music
a. Choices among scales
b. Choices among metres
c. Choice of harmony

1. Antecedently known	Probably in part (e.g. scales).
2. Mentally represented	Yes
3. Systematic	In part: there is apparently systematic and non-systematic variation.
4. Dependent on the input	Yes
5. Deterministic	Probably not but ontogenetic evidence is lacking

Table 3 - Parametric Variation in music?

As before such phenomena as occurring in a critical period are compatible with PV, but are not evidence for it. As with moral judgement, the putative unification of typology and acquisition is moot, as acquisitional evidence for the latter is slim, with much of the literature on music restricting itself to typology. It is also relevant that with harmony (which is of recent development) the limitations of the human voice mean that it can be perceptually triggered but not produced by the learner.

4 Generalisation of Parametric Variation in the domain of animal cognition?

The attempt to generalise PV to other domains of human cognition such as music is moot. Is it more plausible in the domain of animal cognition, in particular birdsong in oscine birds? We begin as before by specifying some of the background differences between the two domains. It is necessary to differentiate birdsong, bird calls and bird mimicry, as PV is of potential relevance only to the first of these. The building blocks of birdsong are 'notes', 'syllables', 'phrases', 'motifs', 'types', and 'bouts', arranged in a hierarchical structure (Brenowitz et al, 1997). As Gardner et al (2005:1046) put it: "Canary song is hierarchically structured: short stereotyped syllables are repeated to form phrases, which in turn are arranged to form songs". The 'notes' on the lowest level of the hierarchy may be supplemented by whistles, buzzes and trills. Assuming that PV operates over elements of the appropriate domain, this difference in 'building blocks' is expected. A more salient difference is that there are about 9000 species of bird (including 4000 species of songbird) and "there is enormous between-species variation in song structure, as well as in the characteristics of song acquisition and production" (Bolhuis & Macphail, 2001:429; cf. Suthers, 1999). This is important as birdsong typology is frequently cross-species, but comparability with language demands intra-species treatment.

Humans are an isolated species and no-one would describe our differences from other primates in terms of parametric choices. The various species of birds are related but (generally) do not inter-breed, suggesting that PV is perhaps irrelevant, as 'setting parameters' would be impossible, especially as "each songbird species seems to go about the process of learning to sing in its own way" (Marler 1999:295). One basic difference is seen in the contrast between 'open-ended learners' (such as canaries) and 'age-limited learners' (such as zebra finches). This contrast could not usefully be viewed as parameter setting, but corresponds if anything to the possibility of learning new structures and/or lexical items throughout life.

As striking as the variety of birdsong is the absence in all species of syntax or semantics. Calls may have some minimal content – e.g. indicating predators, and song has territorial and mate-selection functions, but it mainly manifests variety for its own sake. The absence of semantics is reminiscent of music but the putative absence of syntax[10] is a more radical distinction. It is accordingly necessary to specify what is meant by the contentious observation that there are no 'sentences' in animal communication (Marler, 2000:31). It is true that birdsong shows the recombination of learned sequences in many different ways but we agree with Marler that this is not like human syntax and *a fortiori* could not exhibit recursion. (Marler, 2000:39). Finally, and again reminiscent of the observation about a musical lexicon above, it seems that birdsong repertoires are largely memorised rather than 'creative'.

Despite these obvious differences there are also striking commonalities as pointed out by e.g. Kuhl (1999) and Doupe & Kuhl (1999). Thus, birdsong is parallel to language in being internalist, unconscious, reliant on auditory input, acquired in a sensitive period, and universal (for particular species).[11] Like human infants, birds are sensitive to their own 'language': "young birds must hear the songs of their own species in order to learn them, but when faced with a potentially confusing array of songs, they are able to select the ones of their own species to serve as learning templates" (Whaling, 2000:69). This is comparable to the sensitivity to their own language from intra-uterine experience that new-born infants manifest but is presumably prior to any setting of parameters.

---

[10] When birdsong specialists talk of 'syntax' they are referring to what linguists would call 'phonotactics' – the syntax of phonology. This terminological confusion bears closer scrutiny. Marler makes a distinction between 'Phonocoding' - "the ability to create new sound patterns by recombination simply to generate signal diversity" and 'lexicoding' - "when meaningful elements are syntactically joined" (Marler 2000:31). The latter is supposedly unique to humans. The reference to 'meaning' suggests that what is intended is really semantics rather than syntax: birdsong might well have some syntax but with no semantic function. It is significant that when Suthers (1999:51) observes that brown thrashers probably have in excess of 1000 syllables in their repertoire, he refers to 'syllables' – a phonological construct rather than a syntactic one.

[11] Though often only the male sings.

In later development songbirds appear to go through comparable stages to language-learning infants. Thus 'template memorisation' and 'vocalisation matching' correspond to the child's acquisition of the phonological representations of lexical items of its language and the subsequent mastery of their production. Sub-song; plastic song, full (crystallised) song (cf. Bolhuis, 2005), correspond to babbling, and early and late phonological mastery. There are even closer parallels, of which we mention three.

First, Kuhl observes that in humans "language input alters the brain's processing of the signal, resulting in the creation of complex mental maps" (Kuhl, 1999:424). This is an example of the perceptual magnet effect and the native language magnet whereby infants lose some of their innate discriminatory abilities as a function of exposure to a specific ambient language. A parallel to this magnet effect is found in birdsong. Gardner et al (2005) report on canaries' ability to imitate ill-formed song when young, and how this ill-formed song is then "reprogrammed to form typical canary phrasing".

Second, is the issue of 'selection' versus 'instruction'. Just as with language acquisition there are "genetic contributions to the development of *learned* behaviors" (Marler, 1999:311) and oscine birds show interesting parallels with the 'selection' account of acquisition (Piattelli-Palmarini, 1989). "Songbirds actually inherit much of the information required to generate a normal species-specific vocal repertoire … as though memorization is based not on instruction … but on selective processing, imposed on a fund of *innate* knowledge that is to some degree unique to each species" (Marler, 1999:315).

Third is the putative ability of starlings to acquire recursive grammars. The claim (Gentner et al, 2006; Marcus, 2006; cf. Jackendoff et al, 2006) is that European starlings can be trained to acquire complex recursive grammars and that this "challenge[s] the recent claim that recursion forms the computational core of a uniquely human narrow faculty for language (FLN)" (Gentner et al, 2006:1206). We are convinced by Jackendoff et al's (2006) rebuttal of this claim, casting doubt on both the substance and the implications of the research.

A final parallel between birdsong and language is that birds manifest two kinds of variation: regional (dialectal) and individual (idiolectal); that is, there are birdsong dialects (see e.g. Catchpole, 1991). Searcy & Nowicki (1999) report that, as measured by courtship display to recordings, sparrows are sensitive to differences to the song of conspecifics from New York and Pennsylvania. They emphasise that "geographical variation is the product of chance historical factors" (ibid p.591), and that whereas individual variation is adaptive, geographical variation is not. Presumably the development of dialects is crucially dependent on the selection of particular patterns in the input in language as in birdsong.

The crucial consideration for us is whether any of these characteristics satisfy the criteria for PV. Many, such as occurring in a critical period, having hierarchical structure, illustrating 'magnet' effects, and so on are compatible with PV, but are

not evidence for it. Liu et al (2004) demonstrate individual differences in song learning by zebra finches, including cases where variation was shown among siblings who were "members of the same clutch, and they all imitated the same model" (2004:18178). Crucially, however, the juveniles gradually converge (2004:18180) on the 'same' adult song. This kind of individual variation looks comparable to the individual variation found in consonant harmony in the acquisition of phonology, and the selective choice characteristic of PV plays no role (cf. Doupe & Solis, 1999).

In Table 4 we set out the usual contrasts.

1. Antecedently known	Yes, for some species, but imitation + improvisation suffice without PV.
2. Mentally represented	Yes.
3. Systematicity	Mixed. Gardner et al (2005) report on the perceptual magnet effect and "reprogramming" by canaries. Importantly, however, the reprogramming "occurred in the absence of any exposure to normal canary song", leading to the conclusion that "inferred innate rules forced a complete reprogramming of the imitated song" (2005:1047). None of this seems to be comparable to PV in human learning of phonology.
4. Dependent on the input	We know of no cascade effects in the development of birdsong, but it is clear that many of the details are input-dependent. It is also clear that some individual variation in song production is not parametric but comparable to that found in consonant harmony. The identifiable juvenile nature of the imitations cited by Liu et al (2004) makes it reasonably clear that juvenile birdsong does not always correspond to the adult state.
5. Deterministic	Yes.

Table 4 – Parametric Variation in birdsong?

It is time to take stock. As before, the unification of typology and acquisition is moot, the abilities of starlings are impressive in terms of short-term memory, but the best case for PV may reside in the existence of regional dialects. These dialects appear to be learned rather than inherited (Catchpole, 1991:288; cf. Searcy & Nowicki, 1999) and are clearly functional. However, even here there is no evidence (that we know of) that the choices are antecedently given, and the

complexity of what is learned may not be sufficient to motivate the need for PV: there may be no avian equivalent of Plato's problem.

We began by expressing our scepticism towards Hauser et al's claim that what is unique to human language is recursion. Can we provide an alternative conclusion?

5 Conclusions

We started by envisaging a number of possibilities, summarised below:

(15) a. PV is unique to human language
b. PV is unique to humans but not just to language
c. PV is common to human language and birdsong, but not the rest of cognition
d. PV is common to everything – language, cognition, birdsong …
e. There is no coherent (or uniform) notion of PV.

(For historical reasons we excluded logical possibilities such as (f):

f. PV characterizes e.g. birdsong but **not** human language)

A definitive answer is too much to hope for. However certain tentative observations are in order. Critical periods and universals may both be necessary but are clearly not sufficient for PV. As far as typology is concerned, PV can be adduced harmlessly to describe the limits of variation in all the relevant domains: language, moral judgement, music and birdsong. We view this, however, as somewhat banal. The core interest of PV in linguistics lies in its solution of Plato's problem and the unification of typology and acquisition, so the crucial issue is whether the generalisation also applies in music, moral judgement and birdsong. We suspect that it applies in none – that is, (a) above is correct, but evidence in other domains is lacking. As regards music it is tempting to see a role for PV in the determination of the musical idiom internalised by individuals, but the infantile evidence is lacking and we are concerned that the parallels even then are forced. The reason is that in language (syntactic) PV is limited to the lexicon and there is no comparable construct for music.

So our conclusion is muted. We reject the pessimistic (e) above though on bad days it looks fairly persuasive, and is close to Haider's (1993:1) position that PV is epiphenomenal,[12] and settle on (a). But the reasoning is less decisive than we had hoped, leaving unanswered the question WHY should PV be unique to language? We suspect that the answer is going to be messy and complex. Human language is

[12] We disagree with Haider's claim that "UG cannot have any guiding effect on language acquisition" (1993:4); for discussion cf. Smith, 1990.

remarkably complicated and it is that complexity which makes Plato's problem so hard and so interesting. The other domains despite their interest and richness are not complex in the same way. So we envisage multiple answers: partly (but not exhaustively) recursion; partly, as Jackendoff et al (2006) put it: what is "[u]nique to language is a very large learned vocabulary consisting of long-term memory associations between meanings and structured pronunciations plus varied phrasal syntax". And partly PV, where crucially typology and acquisition are united and the central factor distinguishing language from all the rest is the antecedently available knowledge of the possible choices.

References

Ackema, P., P. Brandt, M. Schoorlemmer & F. Weerman (eds.) (2006), *Arguments and Agreement*. Oxford: Oxford University Press.

Baker, M. (1996) *The polysynthesis Parameter*. New York, OUP.

Bolhuis, J. (2005) "Development of behavior". In J. Bolhuis & L.-A. Giraldeau (eds) *The Behavior of Animals: Mechanisms, Function, and Evolution.* Oxford, Blackwell; pp. 119-145.

Bolhuis, J. & E. Macphail (2001) "A critique of the neuroecology of learning and memory". *Trends in Cognitive Sciences* 5:426-433.

Brenowitz, E., D. Margoliash & K. Nordeen (1997) „An introduction to birdsong and the avian song system". *Journal of Neurobiology* 33:495-500.

Brown, S., B. Merker & N. Wallin (2000) "An introduction to evolutionary musicology". In N. Wallin, B. Merker & S. Brown (eds) *The Origins of Music*. Cambridge MA, MIT Press; pp. 3-24.

Catchpole, C. (1991) "Songs". In M. Brooke & T. Birkhead (eds) *The Cambridge Encyclopaedia of Ornithology*; Cambridge, CUP; pp.283-291.

Chomsky, N. (1971) "On interpreting the world." *Cambridge Review* 92: 77–93.

Chomsky, N. (1975) *Reflections on Language*. New York, Pantheon.

Chomsky, N.(1980) *Rules and Representations.* Oxford, Blackwell

Chomsky, N. (1981a) *Lectures on Government and Binding.* Dordrecht, Foris.

Chomsky, N. (1981b) "Principles and parameters in syntactic theory." In N. Hornstein & D. Lightfoot (eds.), *Explanations in Linguistics.* London, Longman. 123–146.

Chomsky, N. (1986) *Knowledge of Language: Its Nature, Origin and Use*. New York, Praeger.

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

Chomsky, N. (2002) *On Nature and Language*. Cambridge, CUP.

Chomsky, N. (2003) *Power and Terror: Post-9/11 Talks and Interviews*. New York, Seven Stories Press.

Chomsky, N. (2006) *Language and Mind*[3]. Cambridge, CUP.

Doupe, A. & P. Kuhl (1999) "Birdsong and speech: Common themes and mechanisms". *Annual Review of Neuroscience*, 22:567-631.

Doupe, A. & M. Solis (1999) "Song- and order-selective auditory responses emerge in neurons of the songbird anterior forebrain during vocal learning". In M. Hauser & M. Konishi (eds) *The Design of Animal Communication*. Cambridge MA, MIT Press; pp.343-368.

Dresher, E. (1999) "Charting the learning path: Cues to parameter setting". *Linguistic Inquiry* 30:27-67.

Dresher, E. & J. Kaye (1990) "A computational learning model for metrical phonology". *Cognition* 34: 137-195.
Fanselow, G. (1993) "Instead of a preface: Some reflections on parameters"; in G. Fanselow (ed) *The Parametrization of Universal Grammar;* Amsterdam, John Benjamins; pp.vii-xvi.
Fodor, J. (1975) *The Language of Thought.* New York, Crowell.
Gardner,T., F. Naef & F. Nottebohm (2005) "Freedom and rules: The acquisition and reprogramming of a bird's learned song". *Science* 308:1046-1049.
Gentner, T., K. Fenn, D. Margoliash & H. Nusbaum (2006) "Recursive syntactic pattern learning by songbirds". *Nature* 440:1204-1207.
Haider, H. (1993) "Principled variability: Parametrization without parameter fixing". In G. Fanselow (ed) *The Parametrization of Universal Grammar;* Amsterdam, John Benjamins; pp.1-16.
Harris, P. (2000) *The Work of the Imagination.*
Hauser, M. (1997) "Minding the behaviour of deception". In A. Whiten & R. Byrne (eds) *Machiavellian Intelligence II: Extensions and Evaluations*. Cambridge, CUP; pp.112-143.
Hauser, M. (2006) *Moral Minds*. New York; HarperCollins.
Hauser, M., N. Chomsky & W. Tecumseh Fitch (2002) "The faculty of language: What is it, who has it, and how did it evolve?" *Science* 298:1569-1579 (November 22).
Hauser, M., F. Cushman, L. Young, R. K.-X. Jin & J. Mikhail (2007) "A dissociation between moral judgments and justifications". *Mind & Language* 22:1-21.
Hjelmslev, L. (1961) *Prolegomena to a Theory of Language.* Madison, University of Wisconsin Press.
Hockett, C. (1958) *A Course in Modern Linguistics.*
Hyams, N. (1987) "The theory of parameters and syntactic development". In T. Roeper & E. Williams (eds) *Parameter Setting*; Dordrecht, Reidel; pp.1-22.
Jackendoff, R. & F. Lerdahl (2006) "The capacity for music: What is it, and what's special about it?" *Cognition* 100:33-72.
Jackendoff, R., M. Liberman, G. Pullum & B. Scholz (2006) "Starling study: Recursion" http://linguistlist.org/issues/17/17-1528.html.
Jones, D. (2007) "The depths of disgust". *Nature* June 14, 2007; pp.768-771.
Joos, M. (ed.) (1957) *Readings in Linguistics.* Chicago, University of Chicago Press.
Jusczyk, P. (1997) *The Discovery of Spoken Language.* Cambridge MA, MIT Press.
Kayne, R.(2005) "Some notes on comparative syntax, with special reference to English and French". In G. Cinque & R. Kayne (eds) *The Oxford Handbook of Comparative Syntax,* Oxford, OUP; pp. 3-69.
Kuhl, P. (1999) "Speech, language and the brain: Innate preparations for learning". In M. Hauser & M. Konishi (eds) *The Design of Animal Communication*. Cambridge MA, MIT Press; pp.419-450.
Lakoff, G. (1996) *Moral Politics*. Chicago; University of Chicago Press.
Liu, W., T. Gardner & F. Nottebohm (2004) "Juvenile zebra finches can use multiple strategies to learn the same song". *PNAS* 101:18177-18182.
Manzini, M.R. & L. Savoia (2007) *A Unification of Morphology and Syntax.* London, Routledge.
Marcus, G. (2006) "Startling starlings". *Nature* 440:1117-1118.
Marler, P.(1999) "On innateness: Are sparrow songs 'learned' or 'innate'?" In M. Hauser & M. Konishi (eds) *The Design of Animal Communication*. Cambridge MA, MIT Press; pp.293-318.
Marler, P. (2000) "Origins of music and speech: Insights from animals". In N. Wallin, B. Merker & S. Brown (eds) *The Origins of Music*. Cambridge MA, MIT Press; pp. 31-48.
Mikhail, J. (2007) "Universal moral grammar: theory, evidence and the future". *Trends in Cognitive Sciences* 11:143-152.

Mundinger, P. (1999) "Genetics of canary song learning: Innate mechanisms and other neurobiological considerations". In M. Hauser & M. Konishi (eds) *The Design of Animal Communication*. Cambridge MA, MIT Press; pp.369-389.
Muysken, P. (in press) *Functional Categories*. Cambridge, CUP.
Nettl, B. (2000) "An ethnomusicologist contemplates universals in musical sound and musical culture". In N. Wallin, B. Merker & S. Brown (eds) *The Origins of Music*. Cambridge MA, MIT Press; pp. 463-472.
Nettl, B. (2005) *The Study of Ethnomusicology: Thirty-one Issues and Concepts*. Urbana IL, University of Illinois Press.
Peretz, I. (2006) "The nature of music from a biological perspective." *Cognition* 100: 1-32.
Peretz, I. & M. Coltheart (2003) "Modularity of music processing". *Nature Neuroscience* 6:688-691.
Piattelli-Palmarini, M. (1989) "Evolution, selection and cognition: from learning to parameter setting in biology and in the study of language." *Cognition* 31: 1–44.
Pinker, S. (1984) *Language Learnability and Language Learning*. Cambridge MA Harvard University Press.
Postma, G. & J. Rooryck (2007) "Phase-recursion, restricted linguistic systems, and full language". MS Leiden.
Roberts, I. (1997) *Comparative Syntax*. London, Arnold.
Roberts, I. & A. Roussou (2003) *Syntactic Change: A Minimalist Approach to Grammaticalization*. Cambridge, CUP.
Safir, K. (1987) "Comments on Wexler & Manzini". In T. Roeper & E. Williams (eds) *Parameter Setting*; Dordrecht, Reidel; pp.77-89.
Saygin, A., F. Dick & E. Bates (2005) "An on-line task for contrasting auditory processing in the verbal and nonverbal domains and norms for younger and older adults". *Behavior Research Methods* 37:99-110.
Searcy, W. & S. Nowicki (1999) "Functions of song variation in song sparrows". In M. Hauser & M. Konishi (eds) *The Design of Animal Communication*. Cambridge MA, MIT Press; pp.577-595.
Smith, N.V. (1973) *The Acquisition of Phonology: A Case Study*. Cambridge, CUP.
Smith, N.V. (1990) "Can pragmatic fix parameters?" In I. Roca (ed) *Logical Issues in Language Acquisition* Foris, pp. 277-289.
Smith, N.V. (1999) "The structure of noise". *Glot International* 4,5:9. (Reprinted in Smith, 2002).
Smith, N.V. (2002) *Language, Bananas and Bonobos: Linguistic Problems, Puzzles and Polemics*. Oxford, Blackwell.
Smith, N.V. (2004) *Chomsky: Ideas and Ideals* (2nd edition); Cambridge, CUP.
Smith, N.V. (2005) *Language, Frogs and Savants: More Linguistic Problems, Puzzles and Polemics*. Oxford, Blackwell.
Smith, N.V. (2007) "Unification and the unity of Chomsky's thought". *Chomskyan Studies* 2:1-21.
Smith, N.V. & A. Cormack (2002) "Parametric poverty", *Glot International* 6:285-287. Reprinted in Smith (2005).
Suthers, R. (1999) "The motor basis of vocal performance in songbirds". In M. Hauser & M. Konishi (eds) *The Design of Animal Communication*. Cambridge MA, MIT Press; pp.37-62.
Tomalin, M. (2007) "Reconsidering recursion in syntactic theory". *Lingua* 117:1784-1800.
Trehub, S. (2000) "Human processing predispositions and musical universals". In N. Wallin, B. Merker & S. Brown (eds) *The Origins of Music*. Cambridge MA, MIT Press; pp. 427-448.
Trehub, S. & E. Hannon (2006) "Infant music perception: Domain-general or domain-specific mechanisms?" *Cognition* 100:73-99.

Whaling, C. (2000) "What's behind a song? The neural basis of song learning in birds". In N. Wallin, B. Merker & S. Brown (eds) *The Origins of Music*. Cambridge MA, MIT Press; pp. 65-76.

Yip, M. (2003) "Casting doubt on the Onset-Rime distinction". *Lingua* 113:779-816.