

The approach to generative grammar originating with Chomsky (1957) has been enormously successful within linguistics. Seeing such success, one wonders whether a similar approach might help us understand other human domains besides language. One such domain is morality. Could there be universal generative *moral* grammar? More specifically, might it be useful to moral theory to develop an explicit generative account of parts of particular moralities in the way it has proved useful to linguistics to produce generative grammars for parts of particular languages? Should moral theorists attempt to develop a theory of moral universals that is analogous to the theory of universal grammar in linguistics? Can moral theorists develop a “principles and parameters” account of possible moralities inspired by the principles and parameters approach to language in current linguistics? Could there be a “minimalist” program for moral theory inspired by the minimalist program in linguistics?

There will be two main parts to our discussion of these questions. The first will focus on the analogy between generative grammar and moral theory. The second will focus on analogies between universal grammar and theories of moral universals.

In the first part, we give some background and indicate how we are going to understand morality and moral theory. We describe certain aspects of generative grammar and how claims about generative grammars are tested, allowing for a distinction between “competence” and “performance”. We then try to say what a corresponding “generative moral grammar” would be and how it would be tested. We next discuss a number of objections to the analogy between moral theory and generative grammar and indicate possible responses.

In the second part, we discuss certain universal constraints on grammars and consider whether there might be similar constraints on moralities. Then we discuss how linguists describe core aspects of languages in terms of principles and parameters and consider what aspects of moralities might be described in similar terms. After that we make some brief remarks about minimalism.

A brief warning: unlike many other authors (e.g., Mikhail 2000, 2007; Prinz 2008; Sripada 2008, Dupoux and Jacob 2007), we will not argue for, or against, the analogy between linguistics and morality. Our aim, first and foremost, is to offer useful clarifications and distinctions, and to address certain important objections. In doing this, we hope to show that pursuing the linguistic analogy has the potential to yield a fruitful research program, for both philosophers and psychologists, and to give a number of suggestions about promising lines of research. We stress that our claims are only about the *potential* of these sorts of approaches; whether such research will bear fruit in the long run is an empirical question.

Background:

The suggestion of an analogy between grammar and morality has deep roots. In *The Theory of Moral Sentiments*, Adam Smith wrote:

The rules of justice may be compared to the rules of grammar...[they] are precise, accurate, and indispensable...A man may learn to write grammatically by rule, with the most absolute infallibility; and so, perhaps, he may be taught to act justly (Smith 1817, Ch VI, Of the Sense of Duty)¹.

The idea of an analogy between grammar and morality gained more prominence in this century. In part, this may have been due to the development of linguistics as a mature science, and to the close contact between philosophy and linguistics. For instance, Nozick (1968) attempted to provide the beginnings of an account of “moral complications and moral structures,” later reporting that he had meant his discussion “to begin to uncover the structure underlying our moral judgments” in the way in which Chomsky (1957) had intended to uncover the structure underlying our linguistic judgments (Nozick, 1999, pp. 4-5).

Similarly, Rawls (1971, section 9) proposed that, just as linguists were trying to write generative grammars that would accord with and account for their linguistic intuitions, moral theorists might try to develop generative moral theories that would accord with and account for their moral intuitions.

A second factor may have also contributed to the rising prominence of the analogy between linguistic grammar and morality. Within philosophy, there have been many attempts to formulate acceptable principles to account for very particular moral issues discussed by Foot (1967) and later writers—including problems about runaway trolleys. (For a sample, see e.g. Fischer and Ravizza 1992.) As these attempts have become increasingly sophisticated and technical, philosophers and psychologists have considered whether work in generative grammar might provide a useful way to think about these problems.

Increased interest in the analogy between linguistic and philosophy has led to a number of recent papers. For instance, Dwyer (1999) and Harman (1999) discuss analogies between morality and generative grammar. Mikhail (2000) defends “Rawls’ Linguistic Analogy,” offering perhaps the best and most detailed defense of the idea; Prinz (2008) and Sripada (2008) both criticize the analogy, as does Dupoux and Jacob (2007).. Kamm (2007) can be read as an exercise in working out an explicit account of parts of her own linguistic intuition, assuming that intuitive responses to cases “come from and reveal some underlying psychologically real structure, a structure that was always (unconsciously) part of the thought processes of some people” (p. 8, fn. 4). And Hauser (2006) is an extensive development by a psychologist of a moral theory that makes use of a number of analogies with linguistics.

¹ Thanks to Richard Holton for directing us to this early reference.

Preliminaries

In this essay we discuss some of the possible analogies between language and morality and between linguistics and moral theory.

We take linguistics to be the science of language but note that its primary object of study is not language in the ordinary sense in which English, German, Mohawk, and Chinese are languages. Any ordinary language in this sense has different dialects, which blend into one another in ways that do not correspond to national or geographical boundaries. Linguists sometimes say that “a language [in the ordinary sense] is a dialect with an army and a navy,” the point being that what counts as Dutch rather than Flemish is a socio-political issue, not an issue in a science of language. It may be that any two people speakers of a language have somewhat different dialects, with different vocabularies. So, Chomsky and many other linguists are concerned with language as a property of a given individual, assumed to be specifiable by an abstract internal representation or I-grammar.

In order to consider relevant analogies with linguistics, we will use the term *morality* to refer in the first instance to a particular person’s I-morality, that person’s moral standards, perhaps as captured by some sort of (possibly unconscious) abstract representation. (We allow that a person may have more than one morality in this sense by analogy with the way in which bilinguals have more than one I-language.) We distinguish morality in this sense from the principles or conventions of a social group and from the correct moral principles, if there are any and they differ from a particular person’s moral standards. It is of course an important issue whether this way of thinking of morality is as useful as the corresponding way of thinking of language is useful.

By *moral theory* we will mean in the first instance an account of a particular I-morality or an account of I-moralities in general. As above, this notion of moral theory contrasts with other possible notions, e.g. moral theory as an account of the correct moral principles or as an account of the moral principles accepted by a social group.

We find it useful to divide the linguistic analogy into several parts. This is because, in a research program, it might turn out that only some of these analogies bear fruit. Of course, the different analogies between linguistic and moral theory are related, so any distinction must be somewhat artificial. We, however, paint four distinct analogies:

1. An analogy between the generative structure of linguistic grammar and a proposed “generative” specification of a particular individual’s moral standards.
2. An analogy between the universal aspect of linguistic grammar and a proposed universal morality.
3. An analogy between the idea that there is an innate linguistic faculty and the idea that there is an innate moral faculty.
4. An analogy between the minimalist program in linguistics and a corresponding minimalist program for moral theory.

The first analogy has been influential since the 1960s, whereas the others have only more recently been taken seriously. We will use these four analogies to structure our discussion.

Part One: Generative Moral Grammar

As Chomsky (1965) explains, a “generative” grammar of a language would be an explicit grammar that provides principles that fully and explicitly characterize the expressions of a language without relying on linguistic intuition as traditional grammars do. So, for example, Chomsky (1957) formulates explicit rules to characterize English auxiliary verb phrases, so that *Jack has been running* and *Jack may have run* are explicitly counted as well formed but *Jack is have run* is explicitly counted as not well formed. More generally a generative grammar for a particular version of English consists in a set of recursive rules that “generate” the expressions of that version of English in the sense in which axioms and principles of proof might recursively generate the theorems of some formal system.

The word “generative” is potentially misleading. There is no implication that the principles of a generative grammar are principles that a speaker follows in generating what he or she says. For example, suppose a grammar contains a transformation rule that converts a structure underlying an active sentence like *Mary kissed John* into a structure of the corresponding passive *John was kissed by Mary*. To suppose that there is such a rule of grammar is not to suppose that a speaker, either consciously or unconsciously, first constructs the structure of an active sentence and then converts that structure into the structure of the passive sentence. Of course, the speaker may actually do this, but that is a further issue.

A generative grammar of a particular language does not merely generate or explicitly specify the grammatical sentences of that language. It also specifies various other expressions of the language—words, phrases and clauses. Furthermore, it specifies the linguistic *structure* of expressions of the language and characterizes various categories of its words or other morphemes, indicating the nouns, verbs, adjectives, prepositions, etc., as well as the various phrases of which these words are “heads”, as when a noun is the head of a noun phrase or a preposition the head of a prepositional phrase. A generative grammar might also specify important aspects of the meanings of expressions, depending on how those expressions are structured, aspects of meaning having to do with the scope of quantifiers and other such operators. The ambiguity of certain expressions might be represented in terms of the different ways in which they can be structured in this sense.

Generative grammars also specify aspects of the way sentences are pronounced (for spoken languages—something similar is true of sign languages). So, generative grammars relate pronunciations of expressions to possible interpretations. In this respect they map sound-meaning (or phonetic-semantic) relationships.

Testing grammars

As we mentioned above, Chomsky (1965) and many other linguists hypothesize that a speaker has an internal generative grammar, an “I-grammar” (Chomsky 1988). Linguists test partial grammars against the linguistic intuitions of the person whose internal grammatical representation is being investigated. For example, if such a grammar implies that *John is having run* is well formed, the conflict between that result and one’s linguistic intuition counts against

that particular grammar. Similarly, it counts against a grammar of one's language if it implies that *Bob shaved him* is synonymous with *Bob shaved himself*, given one's linguistic intuition that these are not synonymous.

However, linguists allow that considerations other than grammar can affect linguistic intuitions: memory limitations, parsing strategies, false analogies, etc. Although one is unable to process very long sentences, linguists assume that one's grammar allows for sentences of any finite length. The fact that a "garden path" sentence like *the horse raced past the barn fell* strikes one as having too many words in it may be the result of parsing strategies that sometimes lead one wrong (as one might see by considering *the horse (which was) raced past the barn fell*). One's intuition that the *not unhappy king finally went to bed* is well formed may result from a false analogy (Langendoen and Bever 1973).

Why not count as part of grammar factors like memory limitations, etc.? Answer: it has proved to be useful and theoretically illuminating to distinguish one's I-grammar from such other factors. Of course, one might ask: *why* is it that this distinction is useful and theoretically illuminating? For several reasons, though, we'd like to avoid a discussion of that question. First, the topic has been discussed elsewhere and at length (e.g. Chomsky 1965). Second, for our purposes, we need only point to the fact that it *has* proved to be useful (understanding why it is useful is only a secondary problem), as our point is only that such distinctions would likely arise in any attempt to understand morality by analogy with linguistics. We say more about such distinctions below.

This point leads to a related one. Because the theoretical fruitfulness of a distinction depends on the exact details of the science, it is difficult for anyone—however situated—to predict in advance which distinctions may turn out to be useful. Our discussion of the analogy with morality, especially our discussion of how one might carve off the moral faculty from other domains of mind, must therefore be highly speculative.

The "competence performance distinction"

In this connection, Chomsky (1965) introduces a technical distinction between "competence" and "performance." However, it must be said at once that this terminology can be very misleading. For Chomsky, "competence" designates the presumed internalized grammar and "performance" refers to all the other factors that affect linguistic intuition and the ways in which language is used. This terminology can be misleading if confused with the ordinary meaning of "competence," because linguistic competence in Chomsky's sense is "not a skill, a set of habits, or anything of the sort" (Chomsky 1972). A person's linguistic skills (competences in the ordinary sense) count as part of what Chomsky calls performance, while a performance in the ordinary meaning of this word would be an event of a certain sort rather than a skill.

The fact that linguistic intuitions are influenced both by a person's competence and by his performance complicates the testing of proposals about generative grammar. If an intuition has been generated by performance factors alone, the generative grammar need not predict it. Of course, it is the usual situation in science that evidence can be the result of multiple factors. Indeed, as indicated below in our discussion of principles and parameters, data about the internal grammars of other people can be relevant to the assessment of my internalized grammar, by

being relevant to the assessment of general principles that might apply to the grammars of all (humanly learnable) languages and so relevant to the assessment of my internalized grammar in particular.

Explicit specification of moral standards

Moral grammar

An explicit specification of a particular individual's moral standards that was analogous to a generative grammar would fully and explicitly characterize those standards, perhaps assumed to be internally but unconsciously represented by the person's I-morality. The I-morality would have to "generate" descriptions of acts, character traits, situations, etc., descriptions that might include moral assessments in terms of ought, obligation, justice, fairness, right, wrong, responsibility, excuses, justifications, etc. In other words it would determine action-assessment relationships, character-assessment relationships, situation-assessment relationships, etc., in something like the way in which linguistic grammars determine phonetic-semantic relationships.

Above, we suggested that an I-morality must generate *descriptions* of acts, character traits, etc. This is because the principles of a generative morality can apply only to certain sorts of descriptions. For example, a utilitarian morality applies only to descriptions that indicate relevant utilities. The infliction of meaningless violence must be represented as lowering utility before a utilitarian morality can deem the act impermissible. More generally, a generative morality must at least implicitly specify relevantly structured descriptions to which its principles to apply. We come back to this point near the end of this essay.

Nevertheless, one might wonder why these structured descriptions would be part of the I-morality rather than part of a separate grammar of action that yields structures to which the I-morality applies. To this, we have two responses. First, we don't think this is a substantive issue; it is similar to asking why count phrase structure as part of I-grammar rather than something to which I-grammar applies. Second, remember that the I-morality is an abstract specification of an individual's moral "competence," not a theory of "performance." There is no suggestion that one first determines the structure of the situation and then comes up with a moral assessment of it.

Mikhail (2000, see also 2007) suggests that relevant aspects of structure may be arrived at through consideration of the theory of action structure in Goldman (1970) and legal concepts of battery, end, means, side effect, intention, etc. The resulting structures are to some extent like the phrase structures that figure in linguistic theory.

Proposals about I-moralities are testable against moral intuitions in much the way many moral philosophers already test proposed moral principles against intuitions. As in the linguistic case, we must suppose that other considerations beyond one's assumed I-morality affect moral intuitions. In addition to memory and other processing limitations, moral intuitions might be affected by "heuristics and biases" (Gilovich, Griffin, and Kahneman 2002), prejudices, emotional responses, self-interest, etc. For example, Greene (in press) argues that intuitions that appear to conflict with utilitarianism are due to morally irrelevant emotional factors.

Here again, as in linguistics, we can ask whether it makes sense to postulate a distinction between an individual's explicit morality or moral "competence" in contrast with various other factors that determine the person's intuitions and actions. As we stated above, we think this question can only be answered by doing real scientific work, and seeing what distinctions, given the nitty-gritty details of the inquiry, are useful. In linguistics, the corresponding assumption appears to be theoretically fruitful and illuminating. It is perhaps more controversial that this is so for morality, since, for example, it may be that emotional responses play a key role in moral judgment of a sort they do not play in linguistic intuitions. It would be interesting to see a positive case for thinking that the role of emotions problematizes the distinction between performance and competence.

Objections to the analogy with generative grammar

In this section we briefly discuss objections to the very idea of trying to develop analogies between linguistics and moral theory.

Normativity

In reviewing (Rawls 1971), Nagel (1973) observes that for Rawls, ethics is "a bit like linguistics: ethics explores our moral sense as grammar explores our linguistic competence," to which Nagel adds the following objection in a footnote (p. 221): "This seems to me a false analogy, because the intuitions of native speakers are decisive as regards grammar. Whatever native speakers agree on is English, but whatever ordinary men agree in condemning is not necessarily wrong. Therefore the intrinsic plausibility of an ethical theory can impel a change in our moral intuitions. Nothing corresponds to this in linguistics (pace Rawls' suggestion on p. 49), where the final test of a theory is its ability to explain the data."

Putting aside the claim that "the intuitions of native speakers are decisive as regards grammar," which is oversimplified, Nagel's idea is that we do not think that the principles of grammar underlying our linguistic competence have to be intuitively acceptable. When we discover that an unintuitive principle like the coordinate structure constraint (discussed below) is part of our grammar, we take ourselves to have learned something about grammar and are not led to change the way we talk. On the other hand, we do think that basic moral principles should make sense to us. If we discover that an unintuitive moral principle is part of our morality, for example the principle that the force of our obligation to help someone in need varies inversely with their distance from us (Singer 1982, Unger 1996), we need not conclude that we have learned something about moral obligation. Instead, we try to abandon that principle.

Here are three points in response to Nagel's objection. First, we can distinguish a normative account of what is really right and wrong from a psychological account of intuitive commonsense morality, just as we distinguish scientific physics from folk physics. It can be as interesting to investigate intuitive commonsense morality as it is to study intuitive folk physics.

Second, just as generative grammar can reveal principles that linguists had not previously formulated, a generative account of an individual's moral sense might uncover moral principles not previously figuring in ethical theories, principles that we might on reflection think are quite acceptable. Examples might include various formulations of a principle of Double Effect and

Thomson's Deflection Principle (Foot 1967, Thomson 1978). (These principles are stated below.) Kamm (2007) advocates this sort of search for acceptable principles. In this way, investigation into commonsense morality may inform normative inquiry.

Finally, some theorists have argued that the normative moral "facts" are "response dependent" and are constituted in part by something about human psychology, as it is sometimes thought that facts about color are constituted in part by something about the ways we perceive color (Lewis 1989, Johnston 1989, Smith 1994). In that case, morality might be closer to grammar than to physics.

Conflicting moral principles (and dilemmas)

Another apparent disanalogy between morality and language is that moral principles are not, or do not seem to be, exceptionless in the way that principles of generative syntax are exceptionless. Lying is wrong, but not when the Gestapo ask where Jules is hiding. Killing an innocent person is wrong, but not when it is the unavoidable result of turning a trolley into a side track to avoid killing five people on the main track. Moral principles appear to be default principles, describing considerations that have to be balanced against each other. This may seem to rule out the sort of deductive model characteristic of generative syntax.

The matter is complicated, however, since a generative moral grammar might contain a higher-order principle about how to get from the applicable default principles to a verdict about a (description of) a particular case. This might be a constraint satisfaction principle of some sort (cf. Thagard 1989). Indeed, "economy" principles in contemporary minimalist linguistics (Chomsky 1995) may provide models for what is wanted in moral theory. However we cannot explore that possibility here.

Controversial Intuitions

It might also be objected that, early generative grammar appealed to relatively uncontroversial intuitions of the sort mentioned above, while some of the trolley case intuitions appealed to by philosophers like Foot, Thomson, and Kamm are more controversial. For example, Thomson and Kamm appeal to their intuition that, while it is impermissible to push an innocent person in front of a trolley in order to stop it from hitting five other innocent people, it is permissible to divert the trolley into a loop where it will be stopped by hitting the innocent person before looping around to hit the other five. This "intuition" is much less obvious than the intuition that it is ungrammatical to say *Jack is have run* and grammatical to say *Jack has been running*.

Why is this an objection? Controversy involves multiple individuals, but a generative grammar of the sort we are discussing aims to provide an explicit account of a single individual's morality. So, it may seem that only one person's intuitions can be relevant. It is not an objection to this project that different people may have different moral views!

But the objection is really methodological. As in linguistics, the task of developing an explicit account of (a significant part of) an individual's I-morality is going to require cooperative research by several investigators and may be initially possible only to the extent that these investigators are developing an account of I-morality that they themselves share. To the extent

that such scientific cooperation is needed, controversy about particular cases will get in the way, because researchers will not be able to rely on their own intuitions.

In response, it can be observed that many intuitions are relatively uncontroversial, e.g. the intuition that it is worse to send poisoned food to a population than to refrain from sending food, even if the same number of people will die; the intuition that it is permissible and maybe even morally required to turn a runaway trolley that will otherwise hit five innocent people to a side track, even if there is an innocent person on the side track (and there is no loop back to the five); and the intuition that it is not permissible to cut up a healthy visitor to a hospital in order to distribute the visitor's organs to five needy patients who will die without them.

To be sure, utilitarians like Singer (1982) may resist those “relatively uncontroversial intuitions” for theoretical reasons. (We say more about this below.) More generally it must be acknowledged that moral disagreements about particular cases can arise because people accept different moralities. But similarly, people speak different dialects, in which case their linguistic intuitions can diverge.

We add that moral intuitions can be objects of controversy in at least two distinct respects. First, it might be the case that there is widespread disagreement, but that the different parties are each quite certain they are correct. This might be, for instance, the situation with respect to norms governing homosexuality or abortion. Second, the widespread disagreement may be accompanied by uncertainty. The first sort of disagreement is perhaps more likely to represent a difference in I-morality, whereas the second may simply involve difficulties in assessing hard cases.

Maybe no “moral competence”

Of course, it may be that morality cannot be modeled by a moral grammar. Perhaps morality lacks enough systematicity to be generated by any relatively small set of axioms, or perhaps only *some* of morality can be modeled by a generative system.

A second threat arises by considering the psychology of moral judgments. Perhaps moral competence cannot be disentangled from performance influences: how issues are framed, emotional reactions, biases, etc. Compare the familiar suggestion (e.g. Shafer and Srivastava 1990) that people do not have an internalized subjective probability distribution and that their probability judgments are arrived at on the spot as the result of such performance factors. If this were so, then an attempt to formulate a moral generative grammar would be useless, for we would be unable to isolate the set of moral judgments which the grammar ought to generate. Worries of this sort can be met only by the successful construction of partial moral grammars, just as analogous worries about generative grammar have been largely met.

Part 2: Universal moral grammar

We have so far discussed one idea from linguistics that might be relevant to the theory of morality, namely, generative grammar—an explicit specification of the grammatical properties of expressions of a language. The conception of grammar as generative or explicit in this way

has proved quite productive for the study of language. We have so far been discussing whether it might be equally productive for the study of morality to attempt to produce an explicit generative account of an actual morality or part of one.

We now turn to a related issue about linguistic and moral *universals*. Features of language are universal if they occur in all natural human languages that children can acquire as first languages. *Universal grammar* is concerned in part with such linguistic universals and also with limits on ways in which natural human languages can differ. Our present issue, then, concerns the extent to which it makes sense to pursue a universal moral theory which seeks to describe common features of any moralities children can acquire as first moralities—along with an account of how moralities can differ, an account that might be analogous to the account linguistics offers of differences between possible natural human languages.

At least two distinct sorts of moral principles might claim universality. First, there is a commonsensical view that various familiar moral principles are universal in this sense: “You should not steal.” “It is wrong to murder or harm others.” “You should not tell lies.” “You should do your part in useful group activities.” “You should help those worse off than yourself.” “You should not treat others as you would not want to be treated.” “You should do unto others what you would want them to do unto you.” We have already observed that such possibly universal principles tend to be accepted not as exceptionless, but as default principles of varying strengths. We observe here that there do not seem to be equally well known suggestions about what to do when these principles conflict. We will consider below in what ways, if any, these principles are in fact universal.

Second, there are certain non-obvious moral principles, such as the Double Effect, which we will discuss later. These are principles that philosophers propose in order to account for moral intuitions, and they are often not very familiar.

Universal features of language: constraints on rules

In this section, we provide a brief example of a linguistic universal having to do with constraints on transformations. The important point about these constraints is that they are non-obvious universals. Furthermore, there is reason to think that children are predisposed to acquire these constraints.

To begin, consider several instances of phrase structure rules: an adjective like *green* can combine with a noun like *apple* to form a noun phrase *green apple*; a determiner like *a* or *the* can combine with a noun phrase to form a determiner phrase like *a green apple*; a preposition like *from* can combine with a determiner phrase to form a prepositional phrase like *from a green apple*; and so on. Universal grammar places constraints on the phrase structure rules for a given I-language. One such constraint is that, if there are prepositions and so prepositional phrases in which prepositions appear *before* their objects, then the same must be true of verb phrases and noun phrases: verbs must appear *before* their objects and nouns *before* their objects. On the other hand, if instead of prepositions there are postpositions and so postpositional phrases in which postpositions appear *after* their objects, then the same must be true of other verb phrases and noun phrases: verbs and nouns must appear *after* their objects.

Now, transformations indicate how items in phrase structures can be moved around. For example, *wh*-movement allows the derivation of a relative clause like (1) from something like (2):

- (1) which Bob gave to Bill
- (2) Bob gave *X* to Bill.

Topicalization allows the derivation of (3) from something like (4):

- (3) A book, Bob gave to Bill
- (4) Bob gave a book to Bill.

It was noticed early on that there are constraints of various sorts on these rules. For example, there are seemingly complex constraints on transformations. Chomsky (1964) suggested an “A over A” principle that ruled out moving an item of a given type *A* if the item is contained in a larger phrase that is also of type *A*. That turned out to be overly simple and Ross (1967) proposed to replace it with a number of more specific “island constraints”, where “islands” are structures that items cannot be moved out from. Ross’ constraints include a “coordinate structure constraint” and a “complex NP constraint.”

We describe only the coordinate structure constraint. Coordinate structures are special cases of the *A* over *A* principle in which an item of type *A* immediately contains two or more coordinate items of type *A*. Examples: *a man, a woman, and a child; Jack or Harry*. These are noun phrases immediately containing coordinate noun phrases.

Ross’ coordinate structure constraint prevents transformations from moving any coordinate item out of a coordinate structure and also prevents transformations from moving anything contained in a coordinate item out of that item. Suppose, for example, there is a transformation of topicalization which starts with (5) and moves *a book* to the front to yield (6):

- (5) Bob gave a book to Bill
- (6) A book, Bob gave to Bill.

Ross’ constraint prevents topicalization from moving *a book* from a position in a coordinate structure like (7) to yield (8). Similarly, there is no way to obtain the relative clause (10) from (9).

- (7) Bob gave a record and a book to Bill
- (8) * A book, Bob gave a record and to Bill.
- (9) Bob gave a record and *X* to Bill.
- (10) * which Bob gave a record and to Bill

Observe that it is unclear how a coordinate structure constraint of this sort could be learned by children either from explicit instruction or by trial and error. No one had heard about this constraint before Ross (1967), so no one could have told children about it before that time.

Furthermore, children never seem to make errors that consist in violating the constraint, so they don't seem to acquire it by trial and error learning. They seem to be predisposed toward it.

If so, that means children should be predisposed toward such a constraint no matter what natural language they are exposed to, a conclusion that would support an expectation that the same constraint will apply to movement rules in generative grammars of all human languages. That expectation seems to be correct.

Of course, it is possible to invent and use an artificial language that is not subject to the constraint. But such a language might not be easily acquired as a first language by children. We might expect them to acquire a variant subject to the constraint. Furthermore, we might expect that children not exposed to human language (e.g., because they were deaf) who developed a language among themselves (presumably a sign language) would develop a language that was subject to the constraint. There is some evidence that this expectation is correct. Padden (1988) claims that coordinate clauses in American Sign Language (ASL) adhere to the coordinate structure constraint.

Universal moral constraints: non-obvious moral principles

A number of moral philosophers (Foot 1978, Quinn 1993) have suggested that ordinary moral intuitions obey certain non-obvious rules. For instance, it may be that moral intuitions reflect some version of the Principle of Double Effect. Here is one version of that principle:

Double Effect: It is worse knowingly to harm one person *X* in saving another *Y* if (a) the harm to *X* is intended as part of one's means to saving *Y* than (b) if the harm to *X* is merely a foreseen unintended side-effect of one's attempt to save *Y*.

For another example, Thomson (1978) suggests that some ordinary moral intuitions might reflect a principle of the following sort:

Deflection: It is better to save a person *X* by deflecting a harmful process onto another person *Y* than by originating a process that harms *Y*.

As with Ross' coordinate structure constraint, these principles are not generally recognized and it is therefore unlikely that they could be explicitly taught to children. If children do acquire moralities containing such principles, they must be somehow predisposed to do so. If so, we might expect such principles to be found in all moralities that children naturally acquire.

As in the linguistic case, it will be possible to construct more or less artificial moral systems that do not involve such principles, such as utilitarian moralities. But we would expect that children of utilitarians will not easily and naturally acquire their parents' morality but will acquire a version with such principles. John Stuart Mill (1859, 1863, 1873) might be an example. His version of utilitarianism appears to incorporate various principles that do not fit with the version taught him by his father, James Mill (1828). In particular, John Stuart Mill's views about the importance of personal creativity, self-development, and personal liberty led him to adopt a

distinction between higher and lower qualities of pleasure that conflicts with the purely quantitative conception of amount of pleasure characteristic accepted by Bentham and James Mill.

Familiar moral principles

Of course, principles like Double Effect and Deflection are not the sorts of things people usually think of when they hear the phrase “moral universals.” In itself, that is not a problem—linguistic universals are also a far cry from the sort of things one learns in eight-grade grammar. Nevertheless, the question arises as to whether there are other possible moral universals consisting in familiar principles, such as some version of the Golden Rule, or more particular injunctions against stealing, murder, and the harming of innocent people.

Are these principles (e.g., the Golden Rule, don’t murder, etc.) universal and, if so, why are they universal? There are two tricky issues here. First, one must isolate the exact notion in which such a principle is universal. Here we think it might be useful to consider the principles and parameters approach in linguistics. The second issue surrounds the causal explanation of why a moral principle such as “don’t murder” is universal. In language, there is good reason to locate the explanation in an innate predisposition; in the case of morality, this is much less clear.

We’ll begin by introducing the principles and parameters approach in linguistics.

How languages differ: principles and parameters

Baker (2001) begins a highly readable account of current linguistic theory with the story of eleven Navaho Code Talkers. During World War II Allied forces in the Pacific were in difficulty because Japanese cryptographers were able to decode all their messages. The US Marine Corps solved the problem by using the Code Talkers to transmit and receive military communications in Navaho. Japanese cryptographers were unable to “decode” these messages.

Baker notes that this illustrates two important points about languages. On the one hand, Navaho is sufficiently different from English and Japanese that skilled cryptographers were unable to decode the Navaho messages. On the other hand, Navaho is sufficiently similar to English that the Code Talkers could immediately translate messages from English into Mohawk and from Mohawk back into English. How can we account for both the similarities and the differences? Indeed, how can we account for the fact that a child exposed to a language when young will easily acquire that language, when the best cryptographers are unable to make heads or tails of it??

Part of the answer may be that the syntax of a language is largely determined by the setting of a small number of parameters—a parameter that indicates whether a sentence must have a subject as in English, or not as in Italian; a parameter that indicates whether the head of a phrase precedes its complements, as in English, or follows, as in Japanese; and so on. Baker suggests that it may be possible to give something like a structured periodical table of languages, much as there is a structured periodical table of chemical elements. (So, his title, *The Atoms of Language*.)

A child picks up language from its interactions with others; there is no need for explicit teaching. According to the principles and parameters theory, the child uses its limited experience to set a relatively small number of parameters and acquire the “core” syntax of a language. Other “non-core” aspects of syntax, perhaps involving stylistic variation, are learned as exceptions.

Universal grammar includes an account of principles and parameters. To repeat a point mentioned above, it is possible to devise and use a language that is not in accord with the principles of universal grammar, and people can learn to use that language. But children will not naturally acquire that language in the way they acquire languages that are in accord with universal grammar.

Another aspect of the principles and parameters universal grammar approach in linguistics is that one’s conception of the generative grammar of a given language can be affected by considerations involving other languages and ease of language acquisition. The task of the child learner is easier to the extent that as much as possible of grammar is part of universal grammar and so is built into the child’s innate language faculty. So, for example, instead of detailed rules for noun phrases, verb phrases, adjective phrases, and prepositional phrases, the basic theory of phrase structure might become part of universal grammar (as in “X-bar syntax”), so that the grammar of a particular language like English is limited to such matters as whether the head of a phrase (noun, verb, adjective, or preposition) comes before or after its complement. The grammar of a particular language like English might contain no transformational rules; instead there might be a rule of universal grammar that says something can be moved anywhere, subject to universal constraints (Chomsky 1981, van Riemsdijk and Williams 1986).

Let us now briefly consider how these developments and speculations in linguistics might suggest promising avenues for moral theory.

Are familiar moral principles, e.g. don’t murder and the golden rule, universal?

Are moral commonplaces, such as injunctions against murder or the Golden Rule, universal? Answering this question requires that we get clear on the sense of “universal” in play here.

In one respect, it is trivial that all moralities recognize murder to be wrong, because murder is trivially the wrongful killing of a person. On the other hand, moralities can differ as to who is protected by rules against murder and harm. While some contemporary moralities extend some sort of protection to all people (or even to all animate beings sensitive to pain), many moralities have a more restricted range. So, a more accurate statement of a particular morality’s rule against killing might take the form, “It is wrong to kill those who are *G*,” where it may be difficult to explain in words what people are *G* and where which people count as being *G* may vary from one morality to another.

Similarly, consider:

Golden Rule: Do unto others as you would have them do unto you.

This might be better expressed as “Do unto others who are G' as you would have them do unto you.” Again it may be difficult to explain in words what people are G' , and those who are G' may or may not be the same as those who are G .

Similarly, philosophers have argued that moral principles are necessarily *universalizable* in the sense that they apply to anyone in similar circumstances. Perhaps a better conclusion would be that moral principles must be universalizable with respect to those who are G'' , with different unstated understandings about who are G'' .

What are the prospects of a principles and parameters approach to explaining some moral differences? Moralities differ about abortion and infanticide, about euthanasia, about slavery, about the moral status of women, about the importance of chastity in women and in men, about caste systems, about cannibalism, about eating meat, about how many wives a man can have at the same time, about the relative importance of equality versus liberty, the individual versus the group, about the extent to which one is morally required to help others, about duties to those outside one or another protected group, about the importance of religion (and which religion is important), about the importance of etiquette (and which standards of etiquette), about the relative importance of personal virtues, and so on.

Perhaps, as suggested above, all moralities accept principles of the form: *avoid harm to members of group G , share your resources with members of group F* , etc., where G and F are parameters that vary from one morality to another. That is, a principles and parameters approach might provide a framework within which to think about moral diversity—although, as Prinz (2008) points out, such parameters differ from those envisioned in linguistics by having many possible values. The parameters in linguistics have a small number of values, typically two.

To provide a somewhat different example, Dworkin (1993) suggests that all intuitive moralities take human life to have what he calls a “sacred value” that increases during pregnancy and in childhood, reaching a maximum at the beginning of adulthood, and slowly declining thereafter, so that the death of twenty year old person is worse (“more tragic”) than the death of an infant or someone who has reached old age. However, according to Dworkin, individual moralities differ concerning whether an early embryo or fetus has significant sacred value. Dworkin argues that this difference is what lies behind ordinary (as opposed to theoretical) disagreements about abortion. Perhaps there are different settings of a “sacred value” parameter. In this case there might be only two relevant settings, indicating whether there is significant sacred value in the embryo or fetus immediately after conception.

Finally, one might also see parameters at work in anthropological theory. Certain anthropologists have suggested that there are only a small number of ways of organizing societies, with differences in which ways organize which aspects. Fiske (1991) distinguishes communal sharing, authority ranking, equality matching, and market pricing. Shweder et al. (1997) distinguish the “big three” of morality, autonomy, community, and divinity. In these cases, the relevant “parameters” might determine which models fit which aspects of organization in a given society.

Of course, all of this is highly speculative. We will need to have more explicit generative accounts of a variety of moralities before being able to think about principles and parameters.

Familiar moral principles: might they be innate?

In linguistics, there is reason to think that children are innately endowed with the universal grammar. Assuming there were moral universals which could be described using a principles and parameters framework, would it also follow that there is an innate moral grammar?

One ought to be wary of such an inference. First, because such principles are familiar, they can be taught; thus there is not the same reason to suppose children are predisposed to acquire them as there is for universal principles like Double Effect or constraints on transformations. Furthermore, in the case of language, it is difficult to come up with adequate alternative explanations for the universal features of language. In contrast, there are alternative explanations for the purported universal existence of certain moral injunctions. For instance, it might be thought that a society without moral injunctions against murder would not survive. In other cases, what might appear to be a surprising moral universal may turn out to be an a priori triviality. For instance, rules against stealing exist in every society which recognizes private property; but, of course, any society which recognizes private property must have rules against stealing, because—without them—there would be no such thing as private property!

Questions of innateness, then, may be better illuminated by looking at data on acquisition. How does a child acquire an I-morality? Does the child pick up morality from others in something like the way in which language is picked up? Does morality have to be taught? Can it be taught?

Perhaps all normal children pick up the local morality much as they pick up the local dialect, whether or not they receive explicit instruction in right and wrong. Children with certain brain defects might have trouble acquiring morality, just as children with certain other brain defects have trouble acquiring language.

Prinz (2008) objects that moral conventions can be picked up in the way other conventions are, with no need for moral universals, principles and parameters, and other aspects of moral grammar. Similar objections were prematurely raised against universal grammar in the 1960s (e.g. Harman 1967, later retracted in Harman 1973). Whether such objections are plausible depends on details that go beyond what we can discuss in this essay.

Minimalist Moral Grammar?

We'll end with a few, very speculative thoughts on minimalist grammar. The minimalist program (Chomsky 1995) seeks to derive the principles of universal grammar from minimal assumptions, such as that the rules of grammar are minimally recursive and that grammar must interface with some sort of conceptual system and with some sort of perceptual system.

Earlier, we suggested that an I-morality must include specifications of acts, agents, and situations. This raises an interesting question about whether moral theory and practical reasoning ought to be thought of as distinct domains.

The notion of minimalist grammar deepens this puzzle. Whatever system of representation is used to represent actions, it must make use of recursion and therefore involve the sort of “discrete infinity” found in counting and arithmetic, as well as in human language. Hauser, Chomsky, and Fitch (2002) speculate that the ability to use systems with the sort of “discrete infinity” arising from recursive rules is something that distinguishes people from animals. Human languages have grammars with this property but animal signal systems do not. Similarly, human planning can involve recursion—embedding one plan in another. It might be suggested this most basic aspect of planning is somehow the key to human morality as it has been suggested to be the key to human language. If this is accepted, it is an interesting further question how much of morality is contained in the representational system used for practical reasoning. In a Kantian view, of course, all the essential principles of morality are contained in that system. Thus there is a question about the extent to which recursive representation, language, morality, and practical reasoning are all connected.

Finally, it would seem that an important use of language is its enabling of complex thought. The notion that one's I-morality includes a grammar of action suggests, likewise, an important role for morality: the ability to engage in complex practical thought, and not just to get along with others.

We cannot assess these suggestions, but we think they are worth pursuing.

Bibliography

Baker, M. C. (2001). *The Atoms of Language: The Mind's Hidden Rules of Grammar*. New York: Basic Books.

Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.

Chomsky, N. (1964). *Current Issues in Linguistic Theory*. The Hague: Mouton.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press

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

- Chomsky, N. (1988). *Language and Problems of Knowledge: The Managua Lectures*. Cambridge, MA: MIT Press.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- Dupoux, E. and Jacob, P. (2007). "Universal Moral Grammar: A Critical Appraisal," *Trends in Cognitive Science* 11: 373-378.
- Dworkin, R. (1993). *Life's Dominion: An Argument about Abortion, Euthanasia, and Individual Freedom*. New York: Knopf.
- Dwyer, S. (1999). "Moral Competence." In *Philosophy and Linguistics*, edited by Kumiko Murasugi and Robert Stainton. Boulder, CO: Westview Press: 169-190.
- Fischer, J. M., and Ravizza, M. (1992). *Ethics: Problems and Principles*. New York: Holt, Rinehart and Winston.
- Fiske, A. P. (1991). *Structures of Social Life: The Four Elementary Forms of Human Relations*. New York: Free Press.
- Foot, P., (1967). "The Problem of Abortion and the Doctrine of Double Effect." *Oxford Review* 5, 5-15..
- Gilovich, T., Griffin, D., and Kahneman, D., editors, (2002). *Heuristics and Biases*. Cambridge, England: Cambridge University Press.
- Goldman, A. I. (1970). *A Theory of Human Action*. Englewood Cliffs, N.J.: Prentice-Hall.
- Greene, J. (in press). "The Secret Joke of Kant's Soul," *Moral Psychology, Vol. 3: The Neuroscience of Morality: Emotion, Disease, and Development*, W. Sinnott-Armstrong, Ed., MIT Press, Cambridge, MA.
- Harman, G. (1967). "Psychological Aspects of the theory of syntax," *Journal of Philosophy* 64: 75-87.
- Harman, G. (1973). Review of Noam Chomsky, *Language and Mind*, Revised Edition, in *Language* 49: 453-464.
- Harman, G. (1999). "Moral Philosophy and Linguistics," *Proceedings of the 20th World Congress of Philosophy, Volume 1: Ethics*, Klaus Brinkmann, ed. Bowling Green, Ohio: Philosophy Documentation Center: 107-115.
- Hauser, M. D. (2006). *Moral Minds: How Nature Designed Our Universal Sense of Right and Wrong*. New York: HarperCollins.

- Hauser, M. D.; Chomsky, N.; and Fitch, W. T. (2002). "The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?" *Science* 298: 1569-1579.
- Johnston, M. (1989). "Dispositional Theories of Value," *Proceedings of the Aristotelian Society* Supplementary Volume 63: 139-174.
- Kamm, F. (2007). *Intricate Ethics: Rights, Responsibilities, and Permissible Harm*. New York: Oxford University Press.
- Langendoen, D.T., & Bever, T.G. (1973). "Can a not unhappy man be called a not sad one?" In S.R. Anderson & P. Kiparsky (Eds.), *A festschrift for Morris Halle* (New York: Holt, Rinehart and Winston): 392-409.
- Lewis, D. (1989). "Dispositional Theories of Value," *Proceedings of the Aristotelian Society* Supplementary Volume 63: 113-137.
- Mikhail, J. (2000). *Rawls' Linguistic Analogy: A Study of the "Generative Grammar" Model of Moral Theory Discussed by John Rawls in A Theory of Justice*. PHD dissertation in Philosophy, Cornell University.
- Mikhail, J. (2007). "Universal Moral Grammar: Theory, Evidence, and the Future," *Trends in Cognitive Sciences*, 11: 143-152.
- Mill, J. (1828). *Essays on Government*. J. Innes: London.
- Mill, J. S. (1859). *On Liberty*. London: J. W. Parker.
- Mill, J. S. (1863). *Utilitarianism*. London: Parker, Son and Bourn.
- Mill, J. S. (1873). *Autobiography*. London: Longmans, Green, Reader, and Dyer.
- Nagel, T. (1973). "Rawls on Justice," *Philosophical Review* 82, 220-234.
- Nozick, R. (1968). "Moral Complications and Moral Structures," *Natural Law Forum* 13 : 1-50.
- Padden, C. A. (1988). "Grammatical Theory and Signed Languages." In F. Newmeyer (Ed.), *Linguistics: The Cambridge Survey* (Vol. II,). Cambridge, MA: Cambridge University Press.
- Prinz, J. (2008). "Is Morality Innate?" In (Sinnott-Armstrong 2008): 367-406.
- Quinn, W. (1993). *Morality and Action*. Cambridge: Cambridge University Press.
- Rawls, J. (1971). *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Ross, J. R. (1967). *Constraints on Variables in Syntax*. Cambridge, MA: MIT Ph.D. Dissertation. (Published as *Infinite Syntax*, Ablex, 1986.)

- Shafer, G. and Srivastava, R. P. (1990). "The bayesian and belief-function formalisms: A general perspective for auditing." *Auditing: A Journal of Practice and Theory* 9 (Supplement): 110-48.
- Singer, P. (1982). "Famine, Affluence, and Morality," *Philosophy and Public Affairs*, 1: 229-243.
- Sinnott-Armstrong, W., ed. (2008). *Moral Psychology, Volume 1, The Evolution of Morality: Adaptations and Innateness*. Cambridge, MA: MIT Press.
- Smith, A. (1817). *The Theory of Moral Sentiments*. Philadelphia: Anthony Finley.
- Smith, M. (1994). *The Moral Problem*, Blackwell, Oxford.
- Sripada, C. S. (2008). "Nativism and Moral Psychology: Three Models of the Innate Structure That Shapes the Contents of Moral Norms." In (Sinnott-Armstrong 2008): 319-343.
- Thagard, P. (1989). "Explanatory Coherence," *Behavioral and Brain Science* (1989) 12: 435-467.
- Thomson, J. J. (1978). "Killing, Letting Die, and the Trolley Problem," *Monist* 59: 204-217.
- Unger, P. (1996). *Living High and Letting Die*. New York: Oxford University Press.
- van Riemsdijk, H., and Williams, E., (1986) *Introduction to the Theory of Grammar*. Cambridge, MA: MIT Press.