Triggers, Minimalism, and Learnability

Norio Suzuki

Some recent learnability proposals within the Principles and Parameters framework (Chomsky 1981) have held that learnability of a grammar in the limit can be ensured (see the local maxima problem of Gibson and Wexler 1994, though) in terms of extensional considerations (roughly, based on the notion of E-language in the sense of Chomsky 1986; Gibson and Wexler 1994, Berwick and Niyogi 1996, Frank and Kapur 1996), while others have explicitly claimed that intensional resources (I-language constructs) must be available to the learner for the resolution of this problem (Fodor 1998, Dresher 1999, Lightfoot 1999). Berwick and Niyogi (1996) and Frank and Kapur (1996) are basically attempts to refine the framework of Gibson and Wexler (1994), though still in extensional terms: the former make use of Markov chains in modeling the behavior of the Triggering Learning Algorithm (TLA) of Gibson and Wexler (1994), thus arriving at a more precise probabilistic analysis of learnability in parameter spaces, whereas the latter try to formulate a number of conceptions of triggers, weighing the importance of their presence in characterizing convergence of a simple learning algorithm.

Although differing from each other in a number of respects (as we will see below), Fodor (1998), Dresher (1999), and Lightfoot (1999) all essentially agree that input sentences alone cannot ensure the

attainability of the target grammar, some form and amount of I-language ('treelets' as triggers for Fodor 1998, and 'cues' for Dresher 1999 and Lightfoot 1999) being indispensable for the learner to solve the learning problem. Specifically focusing on the presence of ambiguous exemplars in the input for the TLA (arising from the assumption of the impoverished contents of triggers that are represented as sentence types expressed in terms of a set of vocabulary containing only such terms as S, O, V, etc.), Fodor (1998) concludes that the TLA is only a guessing system that may never attain the target grammar under some reasonable assumptions. The rest of Fodor (1998) is an attempt to motivate a conception of trigger that differs from Gibson and Wexler's (1994) in crucial respects and allows the learner to detect parametric ambiguity. Dresher (1999) assumes that grammar acquisition proceeds by means of an ordered path, building on the cue-based approach to parameter setting of Dresher and Kaye (1990). Largely rejecting input-matching models of language acquisition and claiming that input from the target language is only used by the learner as evidence for parameter setting, but crucially not for the purpose of matching target forms, Dresher (1999) follows Dresher and Kaye (1990) in positing that UG associates every parameter with a cue. And one of the two fundamental problems Dresher (1999) notes at the outset is the Epistemological Problem, which deals with the gaps and differences between the vocabulary in terms of which parameters are expressed and the learner's analysis of the input (I will return to this below). One chapter of Lightfoot (1999) is devoted to the explication and illustration of the cue-based model of language learnability in the sense of Dresher and Kaye (1990) and, more recently, of Dresher (1999), and he goes on to amend their view slightly, observing that cues which are realized only in

certain grammars constitute the parameters (p.149). Assuming also the degree-0 learnability of Lightfoot (1991, 1994) and criticizing input-matching models, Lightfoot claims that the learner's triggers should be regarded as a set of abstract structures manifested in the mental representations that result from parsing utterances, the success of the grammar being in no way dependent on the set of sentences that it generates.

One of the principal purposes of this paper is to reassess the viability of sentence *patterns* as TLA triggers (Gibson and Wexler 1994), taking into consideration refinement attempts of the TLA framework made in Berwick and Niyogi (1996) and Frank and Kapur (1996). I also discuss degrees in abstractness of the I-language fragments used in various attempts to overcome certain shortcomings inherent in E-language-based approaches to language learnability in light of recent developments in Minimalism (Fodor 1998, Dresher 1999, Lightfoot 1999).

1 Some General Remarks

In order to solve the Epistemological Problem mentioned above in the sense of Dresher (1999), I claim, the notion of *bootstrapping* will have a certain amount of plausibility and validity in making this problem tractable. Bootstrapping mechanisms come in a variety of forms: phonological, syntactic, semantic, and so forth (Weissenborn and Höhle eds. 2001). Note that it has been observed in the literature that some parameters are set very early in development (e.g., the head-complement parameter), while others slot into place late (e.g., the well-known delay of part of Principle B of the binding theory). Following Bennis, Everaert, and Reuland (eds.; 2000) and Reuland (2001), I assume

that there is division of labor between the different components of the human language system (i.e., C_{HL} (narrow syntax), interpretive procedures, and access to the discourse storage) and that "processes within C_{HL} proper are less costly than processes involving the C-I interface, and these in turn are less costly than processes involving the discourse storage" (Reuland 2001:472). Here we see a ranking of various operations related to the three components in terms of economy considerations, and I assume that this ranking must also be relevant to acquisition. That is, C_{HL}-related operations should be easiest to learn and they should be acquired earliest and most rapidly (presumably, most or all of C_{HL}-related things are innate, apart from the parameter setting portion), interface operations coming next and operations connected with the discourse storage being the last to be acquired. And parameters and parts of parameters are assumed to be scattered over the three components, so the timing of the setting of a given parameter or of part of a given parameter must be a function of the place where it happens to find itself in the three components ; hence, the precedence of the head-complement parameter over (part of) Principle B of the binding theory (actually, prosodic cues may be involved in setting of the former parameter (Guasti, Nespor, Christophe, and van Ooyen 2001); also see below, where it is shown that parameters are largely centered around narrow syntax and that it is, rather, bootstrapping mechanisms that may be situated in different places, roughly in the phonology and semantics). As is clear from this, one and the same parameter may range over more than one component, as in the case of Principle B of the binding theory (assuming that it is a single parameter; alternatively, it may well be the case that it is only a descriptive term as in the case of Passive and thus that parameters may not range

over more than one component). Look at further examples as follows: interpretable/uninterpretable dichotomies of formal features as C_{HL} parameters (as in recent Minimalism) and the setting of Principle A of the binding theory as arising only from narrow syntax resources (specifically, the setting of the governing category parameter in GB terms, and see also the reflexivity framework, where part of what used to be dealt with syntactically is handled in terms of the notion of logophoricity), the V2 phenomenon as an A-P interface parameter (see Chomsky's 1995 approach to it as a PF-phenomenon; notice that this can be stated as an instance of functional parametrization in the parametric system of Roberts and Roussou 2002), and Grodzinsky and Reinhart's (1993) Rule I (Intrasentential coreference), parametrized as a function of variation between child and adult language (pertaining to the discourse storage).[1]

Note that I have assumed so far that lexicons aside, parameters can define variation between languages at the synchronic level (including differences between dialects and idiolects), differences between child and adult language (specifically, during the language acquisition period), and diachronic changes in languages over time (I will propose a more general notion of a parameter, based on a specific interpretation and generalization of the Extended Projection Principle (EPP) later). The above observation concerning parameters might lead to the abandonment of the well-established assumption that parameters are restricted to functional domains of the lexicon (see Chierchia 1998 for a possible semantic parameter and Longobardi 2001 for discussion of this from the viewpoint of functional parametrization; I will return to this in connection with the 'generalized EPP-parametrization,' proposed below).

As for the relationship between parameters and triggers, I assume, partially based on Dresher (1999), that UG associates every parameter with a *bootstrapping* mechanism as its trigger (mainly due to the Epistemological Problem in the sense of Dresher 1999). Bootstrapping mechanisms may range over various components such as phonology, syntax, semantics, etc., perhaps with the following temporal order in acquisition arising from economy and other considerations (Reuland 2001): phonology → syntax → semantics → pragmatics (but see below). Thus, it may be the case that a parameter is set early if it is associated with a phonological bootstrapping mechanism as its trigger (the head-complement parameter), it is set late if it is associated with a semantic/pragmatic bootstrapping mechanism as its trigger, and so forth. This may at least partly account for the acquisition order of the fourteen (English) grammatical morphemes studied by Brown (1973:313-317) and such typical cross-linguistically explored phenomena of early child language as the omission of subjects, the use of non-adult infinitives in main declaratives, and the well-known delay of part of Principle B of the binding theory.

2 Extensional Approaches to Learnability

Notice that Gibson and Wexler's (1994) TLA triggers are presented as sentence patterns expressed in terms of a set of vocabulary including such notions as S, O, V, etc., which, presumably, is based on some arbitrary E-language type of grammar (Chomsky 1986). Since the construct (E-language) is understood independently of the properties of the mind/brain (Chomsky 1986:20), TLA learning procedures must necessarily be ones that belong to the component of as yet obscure general

learning strategies at best (under the assumption that the properties of the mind/brain above stand for domain-specific knowledge). Discussing the 'poverty of stimulus' problem concerning language acquisition, Fodor and Crowther (2002) establish that language learnability cannot be ensured without UG, learning strategies such as the Subset Principle (e.g., Berwick 1985, Manzini and Wexler 1987) and the Uniqueness Principle (e.g., Wexler and Culicover 1980, Pinker 1984, Berwick 1985) being able to contribute to acquisition only in conjunction with innate constraints on the class of possible grammars. They note, for example, that what the Subset Principle alone predicts is totally incompatible with what children actually do during the language acquisition period. Then, it follows that under the TLA the learner cannot possibly attain the target grammar in that TLA learning procedures are based on general learning strategies at best, totally lacking the necessary component of UG (a description of the faculty of language (FL) at the initial state of development), which Fodor and Crowther (2002) have established as one of the indispensable ingredients for grammar acquisition. (See Dresher 1999:44-54 for a detailed criticism of the TLA of Gibson and Wexler 1994.)

However, there may be attempts to remedy this situation while still maintaining an extensional approach, Berwick and Niyogi (1996) and Frank and Kapur (1996) being representative examples. Modeling the behavior of the TLA of Gibson and Wexler (1994) as a Markov chain, Berwick and Niyogi (1996) try to describe formally the conditions for learnability in finite parameter spaces, uncover problematic states in addition to the local maxima described by Gibson and Wexler (as states in the parameter space from which it is impossible for the learner to escape in that there are no local triggers for the grammars (i.e., the

states that the learner is in) to get to the target grammar, given only examples from the target grammar), and characterize convergence times for the learning algorithms quantitatively. Berwick and Niyogi show that Markov structures enable them to compute transition probabilities between grammar states (which the TLA cannot) and that Gibson and Wexler's(1994) adoption of the Greediness Constraint[2] and the Single Value Constraint[3] contributes to making the task of acquisition more difficult and less tractable (while giving some alternatives), and discuss some possible problems arising from Gibson and Wexler's (1994) maturational solution to the local maxima problem with regard to identification in the limit and convergence probability. While they emphasize the necessity of a more precise probabilistic analysis of learnability in parameter spaces and claim to have proposed simple incremental algorithms that overcome a number of problems with the TLA, I still would like to take issue with them on the nature of the grammars they adopt and discuss and the resulting parameter spaces (i.e., E-language grammars with notions such as S, O, V, etc., I presume), taking seriously Gibson and Wexler's (1994) observation that "it may be that the parameters identified here are sufficiently different from the parameters in UG that the problem observed here never occurs" (p.438). Indeed, we will see that the situation is quite different with human language parameters (i.e., parameters in UG) when discussing intensional approaches to learnability. Now, Frank and Kapur (1996) explore the importance of triggers (i.e., sentences that "reveal" the settings of parameters of grammatical variation), formalizing a number of conceptions of triggers with the aim of characterizing convergence of a simple learning algorithm by their presence (p.623). Focusing on the "feasibility requirements" (i.e., limited space, limited

time, robustness) that have to be met by parameter-setting algorithms (PSA) if they hope to belong to some cognitively plausible class of concrete parameter-setting algorithms, Frank and Kapur attempt a formulation of triggers on the basis of several classes of parameter spaces: globally triggered, locally triggered, unambiguously locally triggered, and weakly locally triggered (pp.630-643).[4] They show that Gibson and Wexler's (1994) notions of local and global triggers are only sufficient to guarantee convergence of the stochastic PSA,[5] but not the nonstochastic version, and propose alternatives that can overcome some problems connected with the TLA, rightly noting the significance of the issue of lexical acquisition (in addition to much-discussed grammar acquisition) and the problems of bootstrapping (to which I return below). Frank and Kapur's detailed discussion of a large number of parameter spaces under various conditions and of differences between a stochastic and nonstochastic versions of the PSA is highly impressive and instructive in a number of respects, but I still doubt the validity of the ingredients constituting their grammars and parameter spaces (i.e., E-language things) that arise from their adoption of Gibson and Wexler's (1994) learnability framework as it is (only with minor modifications in theoretical terms, *pace* Frank and Kapur 1996).

3 Intensional Approaches and Triggering Bootstrapping

3.1 Fodor 1998, Dresher 1999, and Lightfoot 1999 from the Viewpoint of Minimalism

Chomsky's (1995) Inclusiveness Condition states that no new features are introduced by C_{HL} besides elements already present in the lexical

items selected for the numeration (see also Chomsky 2000), and Reuland (2001) observes that like the mapping to the A-P interface, the mapping to representations at the C-I interface does not obey the Inclusiveness Condition. Note also Chomsky's (2000) strong minimalist thesis (ontological minimalism, in Martin and Uriagereka's 2000 terms) to the effect that "language is an optimal solution to legibility conditions" and his concomitant observation that "suppose that FL satisfying legibility conditions in an optimal way satisfies all other empirical conditions too: acquisition, processing, neurology, language change, and so on" (p.96). It may be highly instructive to review what the above three papers have to say in terms of these (strong) minimalist assumptions as long as they all claim to make use of I-language fragments in dealing with learnability problems.

Some basic observations and assumptions concerning Fodor's (1998) structural triggers learner (STL) are like the following: (a) ambiguous triggers may mislead the learner further away from the target grammar; (b) a deterministic model that can detect parametric ambiguity is needed; (c) triggers are small structural templates (pieces of tree structure, treelets) that are innate, are stored by FL, and constitute the parametric options offered by UG for languages to make use of if they choose to; (d) UG-provided treelets serve both as triggers and as the parameter values triggered, and are adopted into the learner's grammar if input sentences cannot be parsed without them; (e) the notion of a *supergrammar*, such that it consists of the learner's current grammar but includes all of the UG-defined trigger structures; and (f) since parameter values are potential building blocks included among the ingredients of grammars, they can be combined into a supergrammar for parsing with.

Since Fodor (1998) is not sufficiently explicit about what specific form structural templates would take, in what form they would be stored in FL, how the learner could have access to them (see discussion surrounding the necessity of bootstrapping), etc., I start discussion by closely examining what little information she gives with respect to possible forms of triggers.

(1) Mary saw me.

Commenting on differences between the treatment of the sentence (1), which is compatible both with English parameter settings (SV, VO, and −V2) and with German parameter settings (SV, OV, and +V2), in Gibson and Wexler's (1994) learnability framework, and its structural analyses in terms of standard X-bar theory, where the distinction between the English and German cases is crystal clear, Fodor observes that "once triggers are allowed to include structure, ... the fact that the object is to the right of the trace of the verb ... the trigger for the underlying order of verb and object is the underlying order of verb and object," elaborating further on the specifier-head parameter, to the setting of which the subtree with the specifier of IP in relation to its sister I' is relevant, and on the V2 parameter, whose trigger would be a strong C feature that attracts V+I (early Minimalism), and concluding that "each parameter value is thus associated with its own structural (or in the limiting case, featural) signature, that is, with whatever constitutes its essential contribution to the sentence structures that are licensed by grammars which have that value" (pp.15-17).

Recall Pinker's (1984) bootstrapping problem in language acquisition. Nouns, for instance, do not universally sound the same, or occur

in the same position, or share the same type of meaning. How do children use nonsyntactic information (such as phonology, position, or meaning) to arrive at syntactic knowledge? (Bloom 1999:285) This type of problem can be said to be an "access" problem, i.e., a problem of how to account for mechanisms involved in making it possible for the leaner to acquire things to which she has no direct access (which I call the Access Problem henceforth, using this term as well as the term "bootstrapping problem" in subsequent discussion). Notice also that Guasti, Nespor, Christophe, and van Ooyen (2001), for example, discuss a bootstrapping (through prosody) solution to the setting of the head-complement parameter. I assume that their approach to parameter setting is on the right track, and that parameter setting in general may be carried out through some kind of bootstrapping since the vocabulary in terms of which parameters are couched is necessarily abstract, presumably due to the deductive organization of the grammar (see the Epistemological Problem of Dresher 1999). I then show that the Access Problem becomes pertinent when we look closely at Fodor (1998).

As we have seen above, Fodor's(1998) STL contains as triggers such abstract (i.e., difficult of access) structures as "the lexical verb, or finiteness feature, in I not C" (for the −V2 parameter), "the subtree with the specifier of IP in relation to its sister I'" (for the specifier-head parameter), etc., all of which are expressed in terms of various kinds of structural vocabulary which children are assumed initially not to have access to. Sakas and Fodor (2001) observe that during normal sentence comprehension, the parser (which may be innate) takes as input a surface word string and its job is to establish sufficient structure to permit semantic interpretation of the sentence, plausibly involving the

establishment of empty categories and underlying grammatical relations of the kind needed for parameter setting (p.176). And Fodor (1998) goes on to claim that structural triggers can overcome a number of problems with various parsing models because they are ingredients of grammars and ingredients of sentence trees, all the structural triggers in the UG pool being able to be added into the current grammar to create a larger grammar and the input being able to be parsed with that grammar (p.20). Fodor deals with the problem of trigger recognition (a perceptual problem; in addition to the problem of trigger specification, which she assumed is simple, global, and linguistically authentic in her framework) in terms of parsing, but it somehow seems that "learning by parsing" may constitute the (whole) empirical content of parameter setting, the concept of structural trigger being powerful. It should now be quite obvious that Fodor's (1998) STL model suffers from the Access Problem due to the abstract (difficult of access) properties of her structural triggers. In addition, the STL may be pestered with the parsing paradox to the effect that the learner cannot parse the very sentences she should learn from (Sakas and Fodor 2001:177).

Dresher's (1999) treatment of a number of parameters on the basis of Dresher and Kaye's (1990) cue-based learner focuses primarily on phonology, and in spite of Dresher's claim that the fundamental problems (including the Epitemological Problem; p.28) remain the same both in phonology and in syntax, I still see some important differences between phonology and syntax in that the latter is the generative component while the former is an interpretive one, deferring a unified treatment of the two domains until future research. The reader may benefit from Dresher's (1999) detailed discussion and criticism of some

recent alternative approaches to language learnability such as the TLA (Gibson and Wexler 1994), a genetic algorithm (Clark 1990, 1992, Clark and Roberts 1993), etc.

On the basis of the cue-based theory of language acquisition of Dresher and Kaye (1990) and Dresher (1999), Lightfoot (1999) assumes that cues which are realized only in certain grammars constitute the parameters, i.e., the points of variation between grammars, cues being abstract structures and elements of I-language. We can see that Lightfoot (1999) satisfies Fodor and Crowther's requirement that UG guide the learner in her task of language acquisition. Lightfoot continues that the child scans the linguistic environment for designated structures, i.e., cues (e.g., $_{\text{SpecC}}$ [XP] (V2 cue)). But since cues are abstract elements of I-language, we immediately see that the learner encounters the Access Problem on the assumption that she initially does not have direct access to UG elements. Note that Dresher (1999) observes in this connection that cues to parameters become progressively more abstract and grammar-internal, and that acquisition of representations and acquisition of grammar proceed together (p.27). Presumably, bootstrapping mechanisms (which I discuss below) may be cues in the sense of Dresher (1999) at the *initial* stages of acquisition and Lightfoot's (1999) cues are ones at later stages. Still I claim that Lightfoot's (1999) cues are too abstract for the learner to have access to without some bootstrapping mechanism(s) whose descriptive vocabulary she initially can understand providing necessary measures for her to make a leap into UG. Of course, Dresher's observation that cues to parameters become progressively more abstract can be maintained in terms of the incremental learner (Dresher 1999:41-43). Comparing a number of decoding methods (i.e., the learner's methods of decoding

the parametric signatures of sentences), Fodor (2001) discusses two possible implementations of the abstract cue detector, under which Lightfoot's (1999) cue-based approach to acquisition belongs: a version without prior sentence parsing and one where the input is first fully parsed in order to uncover its more abstract derivational properties on the basis of which I-language cues could be identified (Fodor 2001:744-746). Take, for example, the case of the V2 cue: $_{\text{SpecC}}$[XP]. According to Fodor, the former version (the one without prior sentence parsing) will not work, because there is no obvious way for the learner (not knowing the right grammar) to identify an XP (e.g., a DP) in an unstructured word string and even if she could, it surely would be impossible for her to establish that this phrase is in [Spec, C] position rather than in other positions. And Fodor goes on to conclude that the latter version (the one where the input is first fully parsed) is quite unrealistic, because scanning a sentence for the cue (e.g., $_{\text{SpecC}}$[XP]) for a parameter value (e.g., +V2 of the V2 parameter) entails parsing the sentence with that parameter value (i.e., +V2), sentences, however, not normally being able to be parsed with just one parameter value and the learner having to parse the sentence with +V2 together with many combinations of values of the other parameters until she found one that succeeded for the sentence (since she does not yet know what the target parameter values are). It may rather be obvious that the conclusion of the observations above would lead to the situation where the learner's workload explodes.

3.2 Bootstrapping as Trigger

The notion of bootstrapping implies that on the basis of already

existing knowledge and information processing capacities the learner can make use of specific types of information in the linguistic and non-linguistic input in order to determine the language particular regularities which constitute the grammar (and the lexicon) of her native language (Weissenborn and Höhle 2001). Prosodic, lexico-semantic, conceptual, morpho-syntactic, and pragmatic bootstrapping can be distinguished depending on the type of information that the learner makes use of; e.g., "Physical whole objects are canonical individuals" (mapping between nonlinguistic cognition and linguistic semantics, an instance of conceptual bootstrapping into semantics; Bloom 1999:296), and "NPs refer to individuals" (in the form of a universal of language (hence, UG-provided), mapping between syntax and semantics, an instance of semantic bootstrapping into syntax; Bloom 1999:295). Weissenborn and Höhle (2001) note that the central assumption behind the bootstrapping approach is that there is a systematic relationship between properties of the input at one level of representation, which the learner already has access to, and another level of representation, which is arguably too abstract, giving the intensively studied parallelism between prosodic and syntactic structure or between lexico-semantic and syntactic structure (Gleitman 1990, Pinker 1994) as an example (p.vii).

I assume that UG associates every parameter with a bootstrapping mechanism (in the fashion of Dresher and Kaye 1990 and Dresher 1999, albeit differences between cues and bootstrapping mechanisms), arguing that this follows from Chomsky's (2000) strong minimalist thesis as applied to language acquisition and learnability. I take the relevant proposition to be as follows:

(2) FL is an optimal solution to language acquisition/learnability.

I assume that Chomsky's (2000) observation that his strong minimalist thesis may apply to all other empirical domains (acquisition, processing, neurology, language change, and so on) is on the right track, as far as recent developments in computational diachronic linguistics show mainly on the basis of population genetics that the causal force of language change is language acquisition (i.e., an acquisition-based theory of language change; Yang 2002, Niyogi 2002, Niyogi and Berwick 1998, Lightfoot 1999), and Fodor (1998) gives a major significance to parsing in the task of parameter setting/language acquisition (with important differences noted in the text in judgment on the accessibility of triggers between her approach and mine, though). The following is an example of parameter setting through prosodic bootstrapping, which is actually implemented through what is called the Rhythmic Activation Principle:

(3) a. When you hear sequences of (ws)* within an intonational phrase, set the head-complement parameter with the value head-complement.

b. When you hear sequences of (sw)* within an intonational phrase, set the head-complement parameter with the value complement-head.

(Guasti, Nespor, Christophe, and van Ooyen 2001:237)

Metaphorically speaking, the strong minimalist thesis/ontological minimalism (Martin and Uriagereka 2000) would take parameter setting through bootstrapping as trigger to be a guarantee for FL to

become optimal on the part of legibility-relevant A-P and C-I interfaces, parameters lacking accessibility thanks to the deductive system of FL and bootstrapping arguably being readily available to the learner. I would claim that this is much more than metaphor, and it will now be quite obvious that for FL/UG to work some concrete (physical, I claim; see Weissenborn and Höhle's discussion 2001 concerning the modality in which learning takes place, i.e., spoken or signed language (p.vii)) mechanism is indispensable (as a comprehensive system constituting language). Then, parameter setting should be a whole process through which FL/UG properties become arguably progressively available through bootstrapping mechanisms accessible to the learner. Recall that Fodor (1998) assumes that each one of the UG-provided treelets serves both as trigger and as the parameter value triggered, which situation has been seen to cause her approach to suffer from the Access Problem. As for parameter values, my approach also takes them to be expressed mostly in terms of structural terms whose vocabulary derives from generative work in narrow syntax. Roughly, as long as all begins only when C_{HL} (narrow syntax; the generative component) is operative (due to the organization of FL), all bootstrapping (specifically, the kind of bootstrapping pertaining to parameter setting) may be said to be headed for narrow syntax, which is difficult of access as it is. Now, some observation on the inaccessibility of (narrow) syntax may be in order. For a language to be able to express a large number of things in a number of different situations, it must be fairly complicated. To construct a complicated mechanism, its organization must be fairly profound. To maintain such an organization, the system must be fairly abstract. Hence, the inaccessibility of the system. But the actual fact is that some languages are learnable,

which situation functionally explains the existence of bootstrapping mechanisms readily accessible to the learner. It should, however, be quite easy to understand that all this follows naturally from the strong minimalist thesis/ontological minimalism. Notice that the discussion above leads us to assume that it is bootstrapping mechanisms that belong to any of the three components (i.e., narrow syntax, interpretive procedures, and the discourse storage), but not parameters themselves, which all seem to fall under functional parametrization. Hamann (2002) includes the head-complement parameter, the verb-raising parameters (V-to-I, V2), the clitic parameters, and even the pro-drop parameter among the parameters set very early in the acquisition of syntax. It would be good to discover bootstrapping mechanisms for these parameters because at the earliest stages of development bootstrapping from an accessible into inaccessible domains would be remarkable.

4 EPP-Parametrization

I here refer to a "maximally generalized version of the EPP" (see Chomsky 1981, 2000, among others) as EPP-parametrization, which may encompass possibly a large class of parameters associated with the EPP-feature in that it regulates the possibility of the realization of various functional positions (both X^0 and XP; see Roberts and Roussou 2002 and Chomsky 2000 below, respectively) a la Cinque (1999) with (phonological) material or, perhaps, with UG-provided empty categories (or, Agree alone may suffice in the latter case). I assume the existence of covert phrasal movement, following Pesetsky (2000), Nissenbaum (2000), and Chomsky (2001b). Covert phrasal movement

may also fall under EPP-parametrization, the latter then covering head EPP (PF; Roberts and Roussou 2002), overt phrasal EPP (narrow syntax), and covert phrasal EPP (logical syntax). EPP-parametrization is concerned with the problem of (the head of) which functional (and, perhaps, A-bar) positions may be assigned an EPP-feature, those functional positions thus being filled with (phonological) material (either via Merge or via Move), and the amount of (phonological) material to be inserted into various functional positions may be subject to Suzuki's (2001) economy condition to the effect that you must minimize the effects of the EPP-feature in terms of the amount of pied-piping parameter (pp.25, 32). The relevant mechanism may be such that the very realization of a functional position through the insertion of material into that position can serve for the purpose of checking and deleting the uninterpretable EPP-feature assigned to (the head of) that position. As for the case of X^0 positions, while Roberts and Roussou's (2002) framework differs somewhat from the one presupposed here, we can easily accommodate their insight by replacing their diacritic * assigned to a functional feature with the EPP-feature (see below).

4.1 Functional Projections and EPP-Parametrization

While we have recently seen a number of attempts to further clarify the true nature of the EPP and to gradually dispense with it altogether from the grammar (Martin 1999, Epstein and Seely 1999, Bošković 2002), I still assume the existence of the EPP, such that presumably by marking a position that already has a relevant uninterpretable feature in the usual case, the EPP-feature induces insertion of material into the position (Merge) and copying of material from the goal into the local

environment of the probe (Move), generalized pied-piping somehow determining the size of the material to be moved (in the case of Move; specifically focusing here on the size of "phonological material" moved, which may be variable across grammars; Chomsky 2000, Landau 2001, Suzuki 2001). Drawing on Chomsky's (2000) suggestion that the EPP be extended to X^0 cases as well (p.102), I assume that the EPP applies both to XP and X^0 positions. Consider the following transitive clausal structure:

(4) $[_{CP}$ Spec C $[_{TP}$ Spec T $[_{\upsilon P}$ Subj υ-V $[_{VP}$ t_V Obj]]]] (order irrelevant)

The raising of V to the position of the light verb is assumed to be universal. I submit to Cinque (1999) a much more elaborate clausal structure containing a large number of functional categories, restricting myself here to the usual succinct structure (4), which I presume is enough for my discussion. In (4) only the υP portion is universally filled with phonological material (or, perhaps, UG-provided empty categories), it being determined parametrically across grammars whether the remaining four positions (T, [Spec, T], C, [Spec, C]) will be filled or not. Notice that the latter positions all belong to functional projections, which situation allows EPP-parametrization to be discussed to fall under functional parametrization.

Recall that association between relevant positions (e.g., T (probe) and Subj (goal)) is universally formed by Agree, EPP-parametrization accounting for differences in displacement of relevant material among grammars. As for XP positions, Chomsky (2000) observes that the EPP-feature of T might be universal (but see Alexiadou and Anagnostopoulou 1998 for another approach to "subjectless" sentences

in Celtic, Greek, and Romance) and that for v (its outer Spec) and C the EPP-feature varies parametrically among languages and if available is optional (also Chomsky 2001 and Fox 2000 for optionality). As for X^0 positions, it may be interesting to see how Roberts and Roussou (2002) deal with them in terms of their PF-realization approach, which notates as F* a functional feature F that requires a PF-realization, parametrization being seen as the random assignment of the diacritic * to features typically associated with functional heads. Where the diacritic is assigned to a feature, that feature, F*, must have a PF-realization, * being assigned to F in the lexicon (Borer 1984). The approach makes a further distinction between Merge (more economical) and Move (less economical) in the case of F*, in addition to the distinction between F and F* (pp.25-26). There may exist implicational relations between positions in these functional projections with respect to their possibility of (phonological) realization, but this point is beyond the scope of this paper. (I hope for future research in this direction.) It should now be clear that the term "EPP-parametrization" encompasses possibly a large class of parameters involving the EPP-feature in differing positions in a variety of functional projections and connected with the possible realization of (phonological) material (or, perhaps, UG-provided empty categories) in such positions.

4.2 Triggers for EPP-Parameters

From the discussion above, triggers for EPP-parameters must be accessible to the learner. Surface word strings (sentences) may be accessible, but they cannot exhaustively determine underlying structures because of the presence of ambiguous examples ("parametric ambiguity" in the

sense of Fodor 1998). I seek semantic bootstrapping mechanisms for EPP-parameters, drawing primarily on Rosengren (2002). Rosengren (2002) is an attempt to establish that the EPP is a syntactic device in the service of semantics, but I somehow reinterpret Rosengren's work the other way around so as to make use of various accessible interpretational pieces of information that can arguably be gained from parsing/processing on the part of the learner for the purpose of setting various EPP-parameters on the basis of UG-provided associations between semantic bootstrapping mechanisms (derived from accessible interpretational information) and relevant EPP-parameters. According to Rosengren, the EPP options are exploited at the syntax-semantics interface, which situation I assume may correspond to the existence of (a set of) associations formed by UG between semantic bootstrapping mechanisms and relevant EPP-parameters. The following are some association examples:

(5) a. A specific or generic reading of the subject ⟶ [Spec, T]
b. An existential reading by making the existentially bound event variable visible ⟶ [Spec, Fin]

(based on Rosengren 2002:145)

What (5a) says, for instance, is that if the learner sees the relevant reading in the specific lexical item(s) she somehow comes to determine that the lexical item(s) must be in [Spec, T]; i.e., the setting of the relevant parameter to the effect that [Spec, T] is marked by the EPP-feature and thus filled with (presumably) phonological material in the language she has been exposed to. Recall that the EPP-feature only marks a position into which material can be introduced, but does not

say anything about the amount of that material that is displaced, *wh*-movement being a prime example showing variation in the amount of pied-piped material between child and adult English and among various languages (Gavruseva and Thornton 2001, and Suzuki 2001 for the economy condition concerning the effects of the EPP-feature in terms of the amount of pied-piping).

Notes

1. Rule I: Intrasentential Coreference

   NP A cannot corefer with NP B if replacing A with C, C a variable A-bound by B, yields an indistinguishable interpretation. (Grodzinsky and Reinhart 1993:79)

2. The Greediness Constraint

   Upon encountering an input sentence that cannot be analyzed with the current parameter settings (i.e., is ungrammatical), the language learner will adopt a new set of parameter settings only if they allow the unanalyzable input to be syntactically analyzed. (Gibson and Wexler 1994:411)

3. The Single Value Constraint

   Assume that the sequence $\{h_0, h_1, \cdots, h_n\}$ is the successive series of hypotheses proposed by the learner, where h_0 is the initial hypothesis and h_n is the target grammar. Then h_i differs from h_{i-1} by the value of at most one parameter for $i>0$. (Gibson and Wexler 1994:411)

4. Gibson and Wexler (1994) give the following definitions of global trigger and local trigger (p.409), respectively:

   (i) A global trigger for value v of parameter P_i, $P_i(v)$, is a sentence S from the target grammar L such that S is grammatical if and only if the value for P_i is v, no matter what the values for parameters other than P_i are.

   (ii) Given values for all parameters but one, parameter P_i, a local trigger for value v of parameter P_i, $P_i(v)$, is a sentence S from the target grammar L such that S is grammatical if and only if the value for P_i is v.

5. Roughly, a *stochastic* model of learnability would claim that human

learners also learn a probability distribution describing the applicability of the rules in the grammar in addition to a grammar that explains the data and can be used productively. (Bertolo 2001:3)

References

Alexiadou, Artemis, and Elena Anagnostopoulou. 1998. Parametrizing AGR: Word order, V-movement and EPP-checking. *Natural Language & Linguistic Theory* 16:491-539.

Atkinson, Martin. 2001. Learnability and the acquisition of syntax. In *Language acquisition and leanability*, ed. Stefano Bertolo, 15-80. Cambridge: Cambridge University Press.

Bennis, Hans, Martin Everaert, and Eric Reuland, eds. 2000. *Interface strategies.* Amsterdam Royal Academy of Arts and Sciences.

Bertolo, Stefano. 2001. A brief overview of learnability. In *Language acquisition and learnability*, ed. Stefano Bertolo, 1-14. Cambridge: Cambridge University Press.

Berwick, Robert C. 1985. *The acquisition of syntactic knowledge.* Cambridge, Mass.: MIT Press.

Berwick, Robert C., and Partha Niyogi. 1996. Learning from triggers. *Linguistic Inquiry* 27:605-622.

Bloom, Paul. 1999. The role of semantics in solving the bootstrapping problem. In *Language, logic, and concepts,* ed. Ray Jackendoff, Paul Bloom, and Karen Wynn, 285-309. Cambridge, Mass.: MIT Press.

Borer, Hagit. 1984. *Parametric syntax.* Dordrecht: Foris.

Bošković, Željko. 2002. Expletives don't move. In *NELS* 32:21-40. GLSA, University of Massachusetts, Amherst.

Brown, Roger. 1973. *A first language.* Harmondsworth: Penguin Books.

Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural Language Semantics* 6:339-405.

Chomsky, Noam. 1981. *Lectures on government and binding.* Dordrecht: Foris.

Chomsky, Noam. 1986. *Knowledge of language: Its nature, origin, and use.* New York: Praeger.

Chomsky, Noam. 1995. *The minimalist program.* Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: Essays in minimalist syntax in honor of Howard Lasnik,* ed. Roger Martin, David Michaels, and Juan Uriagereka, 89-155. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2001a. Derivation by phase. In *Ken Hale: A life in language,* ed. Michael Kenstowicz, 1-52. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 2001b. Beyond explanatory adequacy. Ms., MIT, Cambridge, Mass.

Cinque, Guglielmo. 1999. *Adverbs and functional heads.* Oxford: Oxford University Press.

Clark, Robin. 1990. *Papers on learnability and natural selection.* Technical Reports in Formal and Computational Linguistics, No.1, Universite de Geneve.

Clark, Robin. 1992. The selection of syntactic knowledge. *Language Acquisition* 2:83-149.

Clark, Robin, and Ian Roberts. 1993. A computational model of language learnability and language change. *Linguistic Inquiry* 24:299-345.

Dresher, Bezalel Elan. 1999. Charting the learning path: Cues to parameter setting. *Linguistic Inquiry* 30:27-67.

Dresher, Bezalel Elan, and Jonathan D. Kaye. 1990. A computational learning model for metrical phonology. *Cognition* 34:137-195.

Epstein, Samuel David, and T. Daniel Seely. 1999. SPEC-ifying the GF "subject;" eliminating A-chains and the EPP within a derivational model. Ms.

Fodor, Janet Dean. 1998. Unambiguous triggers. *Linguistic Inquiry* 29:1-36.

Fodor, Janet Dean. 2001. Setting syntactic parameters. In *The handbook of contemporary syntactic theory,* ed. Mark Baltin and Chris Collins, 730-767. Oxford: Blackwell.

Fodor, Janet Dean, and Carrie Crowther. 2002. Understanding stimulus poverty arguments. *The Linguistic Review* 19:105-145.

Fox, Danny. 2000. *Economy and semantic interpretation.* Cambridge, Mass.: MIT Press.

Frank, Robert, and Shyam Kapur. 1996. On the use of triggers in parameter setting. *Linguistic Inquiry* 27:623-660.

Gavruseva, Elena, and Rosalind Thornton. 2001. Getting it right: Acquisition of *whose*-questions in child English. *Language Acquisition* 9:229-267.

Gerken, LouAnn. 2001. Signal to syntax: Building a bridge. In *Approaches to bootstrapping: Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition,* vol. 1, ed. Jürgen Weissenborn and Barbara Höhle, 147-165. Amsterdam: John Benjamins.

Gibson, Edward, and Kenneth Wexler. 1994. Triggers. *Linguistic Inquiry* 25:407-454.

Gleitman, Lila. 1990. The structural sources of verb meaning. *Language Acquisition* 1:3-55.

Grodzinsky, Yosef, and Tanya Reinhart. 1993. The innateness of binding and coreferenece. *Linguistic Inquiry* 24:69-101.

Guasti, Maria Teresa, Marina Nespor, Anne Christophe, and Brit van Ooyen. 2001. Pre-lexical setting of the head complement parameter through prosody. In *Approaches to bootstrapping: Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition,* vol. 1, ed. Jürgen Weissenborn and Barbara Höhle, 231-248. Amsterdam: John Benjamins.

Hamann, Cornelia. 2002. *From syntax to discourse.* Dordrecht: Kluwer Academic.

Höhle, Barbara, Jürgen Weissenborn, Michaela Schmitz, and Anja Ischebeck. 2001. Discovering word order regularities: The role of prosodic information for early parameter setting. In *Approaches to bootstrapping: Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition,* vol. 1, ed. Jürgen Weissenborn and Barbara Höhle, 249-265. Amsterdam: John Benjamins.

Landau, Idan. 2001. Control and extraposition: The case of super-equi. *Natural Language & Linguistic Theory* 19:109-152.

Lightfoot, David. 1991. *How to set parameters.* Cambridge, Mass.: MIT Press.

Lightfoot, David. 1994. Degree-0 learnability. In *Syntactic theory and first language acquisition: Cross-linguistic perspectives, volume 2: Binding, dependencies, and learnability,* ed. Barbara Lust, Gabriella Hermon, and Jaklin Kornfilt, 453-471. Hillsdale, N.J.: Lawrence Erlbaum.

Lightfoot, David. 1999. *The development of language: Acquisition, change, and evolution.* Oxford: Blackwell.

Longobardi, Guiseppe. 2001. How comparative is semantics? A unified

parametric theory of bare nouns and proper names. *Natural Language Semantics* 9:335-369.

Manzini, M. Rita, and Kenneth Wexler. 1987. Parameters, binding theory, and learnability. *Linguistic Inquiry* 18:413-444.

Martin, Roger. 1999. Case, the Extended Projection Principle, and Minimalism. In *Working Minimalism*, ed. Samuel David Epstein and Norbert Hornstein, 1-25. Cambridge, Mass.: MIT Press.

Martin, Roger, and Juan Uriagereka. 2000. Introduction: Some possible foundations of the Minimalist Program. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik,* ed. Roger Martin, David Michaels, and Juan Uriagereka, 1-29. Cambridge, Mass.: MIT Press.

Nissenbaum, Jonathan, W. 2000. Investigations of covert phrase movement. Doctoral dissertation, MIT, Cambridge, Mass.

Niyogi, Partha. 2002. The computational study of diachronic linguistics. In *Syntactic effects of morphological change,* ed. David W. Lightfoot, 351-365. Oxford: Oxford University Press.

Niyogi, Partha, and Robert C. Berwick. 1998. The logical problem of language change: A case study of European Portuguese. *Syntax* 1:192-205.

Penner, Zvi, Karin Wymann, and Jürgen Weissenborn. 2001. On the prosody-lexicon interface in learning word order: A study of normally developing and language-impaired children. In *Approaches to bootstrapping: Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition,* vol. 1, ed. Jürgen Weissenborn and Barbara Höhle, 267-293. Amsterdam: John Benjamins.

Pesetsky, David. 2000. *Phrasal movement and its kin.* Cambridge, Mass.: MIT Press.

Pinker, Steven. 1984. *Language learnability and language development.* Cambridge, Mass.: Harvard University Press.

Pinker, Steven. 1994. How could a child use verb syntax to learn verb semantics? *Lingua* 92:377-410.

Reinhart, Tanya, and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24:657-720.

Reuland, Eric. 2001. Primitives of binding. *Linguistic Inquiry* 32:439-492.

Roberts, Ian, and Anna Roussou. 2002. The history of the future. In *Syntactic effects of morphological change,* ed. David W. Lightfoot, 23-56. Oxford: Oxford University Press.

Rosengren, Inger. 2002. EPP: A syntactic device in the service of semantics. *Studia Linguistica* 56:145-190.

Sakas, William G., and Janet Dean Fodor. 2001. The structural triggers learner. In *Language acquisition and learnability,* ed. Stefano Bertolo, 172-233. Cambridge: Cambridge University Press.

Suzuki, Norio. 2001. Minimalism and its implications for language acquisition and learnability. In *Kobe Shinwa studies in English linguistics and literature* 21:14-43. Society of English Linguistics and Literature, Kobe Shinwa Women's University, Kobe.

Thornton, Rosalind, and Kenneth Wexler. 1999. *Principle B, VP ellipsis, and interpretation in child grammar.* Cambridge, Mass.: MIT Press.

Weissenborn, Jürgen, and Barbara Höhle. 2001. Introduction. In *Approaches to bootstrapping: Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition,* vol.1, ed. Jürgen Weissenborn and Barbara Höhle, vii-xvii. Amsterdam: John Benjamins.

Wexler, Kenneth, and Peter Culicover. 1980. *Formal principles of language acquisition.* Cambridge, Mass.: MIT Press.

Yang, Charles D. 2002. Grammar competition and language change. In *Syntactic effects of morphological change,* ed. David W. Lightfoot, 367-380. Oxford: Oxford University Press.