Universal bilingualism*

THOMAS ROEPER
University of Massachusetts

Lexically linked domains in language allow a speaker to formulate incompatible rules. How should they be represented theoretically? We argue that a speaker has a set of mini-grammars for different domains so that, in effect, every speaker is bilingual. It is argued that Tense or Agreement Checking, V-2 for quotation, and resumptive pronouns, all lead to bilingual representations. In addition, this perspective on Theoretical Bilingualism suggests that optionality and stages in the acquisition of an initial grammar should also be characterized as a form of bilingualism.

We argue that a narrow kind of bilingualism exists within every language.[1] It is present whenever two properties exist in a language that are not stateable within a single grammar.

We label this claim Theoretical Bilingualism (TB). This view is orthogonal to the obvious social dimensions of bilingualism which understandably have given predominant stature to the sociolinguistic perspective on bilingualism.[2] The social notion of bilingualism – impressive command of two different languages – is very strong. That sense of bilingualism can make it difficult to see that deep theoretical properties of mental structure, apparent in tiny grammatical variations, are also forms of bilingualism.

Much of what we shall claim about multiple grammars has been claimed before. Two features distinguish our approach from previous ones: (i) we use the concept of Theoretical Bilingualism to capture recalcitrant features of first language acquisition, in particular, optionality and lexical variation; and (ii) we utilize Minimalist theory to state in terms of economy where bilingualism within a language is predictable.[3]

The details of bilingual variation are often accurately described as exhibiting a continuum, as one finds for the Romance languages around the Mediterranean. In this essay, I proceed from the assumption that wherever one finds a continuum, or historical gradualism, a more refined level of analysis will reveal discrete phenomena. Thus we aim to identify and dissolve a few of the "continuum" phenomena about bilingualism, while leaving most of the puzzles unadvised.

We begin with a distinction between Language and Grammar from Chomsky (1986). Chomsky distinguishes between Internalized language (= grammar) and Externalized language (=set utterances that can be produced). He argues that E-language may not be ultimately coherent. In discussion he notes:

> we exclude, for example, a speech community of uniform speakers each of whom speaks a mixture of French and Russian (say an idealized version of the 19th century Russian aristocracy). The language of such a speech community would not be "pure" in the relevant sense because it would have "contradictory" choices for certain of these options.

We argue that every language, looked at closely, will involve some domains where "contradictory" choices are made and therefore a hidden bilingualism exists.

* Thanks to Uschi Lakshmanan, Rosemary Tracy and Jürgen Meisel for commentary, to Bart Hollebrandse for discussions, and to several anonymous reviewers for commentary. The essay is written from the perspective of someone who works primarily in first language acquisition. Jürgen Meisel (pc) has helped to bring a broader perspective to the claims made in this essay. He points out, not surprisingly, that the formulation of interpenetration has been an issue for variation theorists for many years, going back to the Junggrammatiker and continuing to the 1970s in the work of C. J. Bailey (1973) and Derek Bickerton (1975).

[1] The concept of bilingualism has never received a widely acknowledged formal definition (to my knowledge). One can even ask: should it receive a clear formal definition? Its cousins – dialects, interlanguage, foreign language, and speech register – all remain important social terms, but unclear theoretical terms. Dialects, for instance, are sometimes defined as "mutually intelligible" languages, which is a valuable human and holistic characterization, but not a formal one.

[2] Power, exclusion, and prejudice all flow from the ability to speak two languages. Power comes from being able to be in two worlds at once. Exclusion comes from the fact that some people can be deprived of important knowledge when others make an effortless shift to an incomprehensible language. Prejudice comes from the seeming imperfections that arise when one language influences another. A mere hint of an accent can seem to the hearer to represent an alien culture. These factors may play a role in motivating people to maintain or avoid bilingualism – even the very narrow sort discussed here – but we shall not address this question.

[3] See Rubin (1996) for a similar discussion of bilingualism as lexical variation.

Address for correspondence
Department of Linguistics, University of Massachusetts, Amherst, MA 01003, USA
E-mail: roeper@linguist.umass.edu

In traditional terminology, both options of mutually exclusive parameters are chosen.

This thesis has implications for two current assumptions in acquisition research (A, B):

(A) The child passes through stages.
(B) Certain rules are optional.

From the TB perspective, a child who is apparently "between stages" is utilizing two (or more) grammars, one of which may eventually disappear. We argue that there is no coherent concept of Stages because separate lexical word-classes may independently use "earlier" or "later" forms of grammar. The result is that incompatible features of grammar may be used by a child simultaneously.

Moreover, under TB, the notion of optionality can be eliminated. If a rule in a child's grammar appears to shift from "optional" to "obligatory", then, in reality, one of two sides of the optionality represents a grammar that has been deleted. We are now purifying the term grammar to include the claim that any consistent grammar cannot have contradictory rules. Therefore one must postulate two grammars, even if they differ only in a single rule.

This is an important step from a formal perspective under what is known as Subset Theory.[4] The logic of learnability theory is that optional rules cannot be eliminated by any straightforward mechanism in the process of acquisition, since no positive input shows that an optional rule is incorrect. In other words, incorrect optional rules create a superset which must be restricted to a subset. No mechanism is available for such a derivation. Movement from a subset to a superset, however, is clearly motivated by input evidence: a new sentence does not fit into the existing grammar, and forces the grammar to be revised. Elimination of optional rules is therefore a step forward in learnability terms, but new questions arise about the relationship among grammars under the assumption that all speakers are bilingual.

A natural extrapolation of this claim is to assert that a person has numerous grammars: every lexical class with rules that are incompatible with another class should constitute a separate grammar. It sounds unwieldy and implausible to argue that a person has a dozen grammars. The essence of this assertion may, nonetheless, be true. It implies that the notion of *a grammar* should change to a more local conception.

One might at this point object that we have not solved linguistic problems but rather turned them upside down. We no longer wonder how and why exceptions exist, since they can all be seen as mini-grammars. Instead, we ask how and why exceptions are eliminated in favor of any far-reaching systematicity in grammar. Indeed, we have traded in one set of problems for their opposites. A shift in perspective, however, can lead to new principles. One claim we shall make is that where two grammars are present, one may represent a *Minimal Default Grammar* definable in terms of economy.[5] Nonetheless, most of the questions about when exceptions survive or disappear remain.

[4] See Berwick (1985).

Universal bilingualism

The notion of Theoretical Bilingualism that we advocate can be defined within the Minimalist Theory of syntax recently presented by Chomsky (1995). We shall provide simply a sketch of that view and concentrate upon some empirical observations.

An example

Let us begin with an example. Children pass through a period in which they will simultaneously say both "I want" and "me want" (or "him want"/"he wants").[6] There are several logical approaches to this phenomenon.

(1) Each form ("I want" and "me want") represents different structures in the same grammar. One might argue that "me" is an emphatic form of "I" (but note that it does not generally receive emphatic stress).[7]
(2) Each form has a different thematic function in a grammar (Budwig, 1989). For instance it has been argued that "me want" is linked to stronger agentive situations.
(3) Each form represents a different stage in child grammar.
(4) Two forms result because Agreement-marking is optional in the child's grammar: "I want" or "he wants" entails Agreement and "me want" does not. The child's grammar changes to make Agreement obligatory.[8]

[5] P. Muysken (pc) has suggested something of this kind to me. See Penner & Wymann (1998) for further discussion of where Minimal Default Grammars function in language acquisition. See also Penner and Roeper (1998).

[6] See Vainikka (1994) for arguments that "me" and "my" are default forms that can appear within VP. It is quite likely that "my" has a distinct analysis from "me", but we will not explore that option in these terms.

[7] See Roeper and De Villiers (1992), Abdul-Kareem (1996) and Schütze and Wexler (1996) for discussion and references for this phenomenon.

[8] Powers (1996) argues that forms like "I want" precede and coexist with the rarer forms "me want". She suggests that there is a chain between an IP subject and a VP subject and "me want" articulates only the VP level, while "I want" reflects a structure

The alternative to all of these approaches is:

(5) Bilingualism: the child has two grammars, one with Agreement and the other without:
G1: Tense Phrase = +/- Tense, +/- Agreement
G2: Tense Phrase = +/- Tense

Roeper and Rohrbacher (1994), based on Speas (1994), argue that UG allows adult grammars that lack AGR, as in G2. Chomsky (1995) argues that AGR is a feature on a Tense Phrase, which makes this scenario even more plausible. It means that a child is simply missing a formal feature, not an entire node.

One possibility is that the English-speaking child abandons G2 (no Agreement), which is socially seen as a pre-school grammar, as it moves into school and toward adulthood. In other words, it is possible that the abandonment of one grammar from a set of grammars could be motivated by social reasons that are external to any particular grammar itself. In that case, the grammar remains but is simply not used. The idea that it continues to be present is suggested by the fact that we can recognize "me want" as child grammar. This "social analysis" is a logical possibility and should remain as an hypothesis.

All of our references to social factors are rudimentary. (One should consult the sociolinguistic literature for more appropriately refined accounts.) In what follows, we will continue to make vague reference to "social factors" as an expression intended to cover a myriad group of factors which may determine the use of grammar but are not expressible in grammatical notation. Careful study of these factors may reveal systematic interfaces where the vocabulary of grammatical notation can be seen as equivalent to other dimensions of cognition. How, for instance, does the cognitive notion of Agent map onto the linguistic notion?

We shall focus on a more tractable possibility: that principles of grammar can eliminate one or another grammar.[9] First we will discuss the role of inference in the use of incomplete grammars.

Interface economy: limiting the role of inference

Adults, like children, are more or less explicit depending upon the social occasion. For instance, if one enters a store and says either (i) "milk" or (ii) "I want milk", both utterances have the same ultimate meaning, but (ii) is clearly more explicit. Situational inference, not written into the grammar, makes (i) just as acceptable. Let us formulate this as a constraint:

Meaning *explicitness* is valued more highly than non-explicitness.

In current terms, if one has two possible *Numerations* (two different selections of items from the lexicon) which define what will be explicit, then the Numeration which leaves less to extra-grammatical inference is preferred. This option is theoretically attractive, but it requires elaboration. In effect, then, we would be elaborating linguistic theory to allow one to prefer one Numeration over another, based on a non-grammatical factor. Therefore it would fall into the domain of *interfaces* between grammar and other cognitive systems. Current models treat different Numerations as simply non-comparable, just as two sentences on different topics are non-comparable. In the example under discussion, G2 is more economical, but less explicit because it contains no AGR node.

It is possible that notions of interface economy, which compare Numerations, will be relevant to the explanation of how a child rejects early grammars, but we will restrict our attention in this essay to the claim that children retain multiple, partial grammars for a single "language".

Economy of representation

It is important to recognize that no regular input justifies the expression "me want", or G2.[10] It is effectively a spontaneous expression derived from innate knowledge of Universal Grammar.[11] What is its status? We will argue below that the two grammars are not equal: G2 follows *economy of representation*. Economy of representation is a relatively new perspective developed by Chomsky (1995) on what constrains possible grammars. In a broad intuitive sense, economy favors less structure and shorter movement rules. We argue that representations like "me want", if economical, can be generated directly from Universal Grammar without an input trigger, under Default Case assignment. Abdul-Kareem

like [IP I_i [VP $pro_{i\ [want]}$]] with a chain between the two subjects. Any analysis must, however, explain why these structures should coexist. No theory of economy will give them equal status. It is inevitable therefore that a concept like bilingualism must be invoked if one wants to leave the concepts of economy within grammar undisturbed.

[9] For instance, the addition of obligatory formal features as these are recognized will change the grammar. See Roeper (1996) for discussion.

[10] Emphatic expressions utilize the Default case and Default tense in English: "Me sing, never!" These could be utilized in the process of identifying the default in English. See Abdul-Kareem (1996) for more argument and evidence that it is question-dialogues which identify the default for the child.

[11] Note that Bickerton (1981) also claims such structures for creole languages.

(1996) shows how elliptical utterances enable a child to identify Default Case.

We have now outlined two criteria that might be relevant in the rejection of a grammar: (i) *economy of representation* and (ii) *meaning explicitness*. As in the "milk" example, how much of one's intention will appear in explicit form and how much left to inference? In formal terms: how extensive will the Numeration be? These two criteria, quite obviously, have opposite characteristics: one favors more, the other less, elaborated structures. We expect the child to go through three stages:

(1) Minimal grammar (me want),
(2) Minimal grammar (me want) and more explicit grammar (I want),
(3) rejection of minimal grammar in favor of more explicit grammar (I want).

Numeration and inference

The selection of a Numeration, in turn, depends in part on a judgment of how much shared inferential information interlocutors have. Here the child may make richer, and partly unwarranted, assumptions. That is, the child assumes a larger shared domain than the adult and fails to communicate adequately. Thus when a child says "that" and the adult responds "do you want something, which thing?" then the child has utilized excessively rich inferences, since the adult must ask for further information.

What does the bilingual speaker do? One might imagine that an insecure bilingual speaker will choose a grammar in terms of context: if the hearer shares context, then a less explicit grammar will work. If one grammar permits subject-drop, and the subject is contextually clear, then this contextual circumstance may influence the choice of grammar. This option may hold for the child bilingual, the adult who controls several dialects, and the true bilingual who selects, say, Spanish or English on different occasions.

Limits to inference

It is important to realize that every grammar does not allow all inferable information to be absent. If the topic of conversation refers to the past, one is not therefore (in Standard English) allowed to delete all references to the past. And although a Noun Phrase may be manifestly singular, it does not entitle one to delete an Agreement marker and say "Mary sing" instead of "Mary sings". Presence of AGR or Tense is immune to available social inferences in Standard English. Once again, we cannot fail to have Agreement -s in "she sings" simply because we derive from context that the verb should be interpreted in the present tense and refer to a singular subject.

So where is inference deemed insufficient by the grammar? When must we use grammar in addition to context? These are very deep questions to which there are no straightforward answers. While we cannot delete a singular Agreement marker in Standard English, we can, when in a context where five people are pushing a car, say "push" instead of "push the car". So context allows the deletion of an entire object, but not the deletion of an Agreement marker.

How is this pertinent to Theoretical Bilingualism? Once again, if one has a choice of languages or dialects, one might decide to choose the dialect which allows the greatest, or least, use of context. In African-American English, for instance, the Agreement and tense markers are generally seen as "deleteable" when context is explicit. In our perspective, AGR and Tense are never deletable, but one can choose a grammatical dialect in which they are not required.

In sum, bilingualism, or code-switching in context, can allow one to evade those features of one grammar immune to contextual information by choosing another grammar where context is utilized. The effect is to shift speech register, since heavy reliance on context conveys informality. All of this is a slightly more formal statement of what is regarded as a common sense view of bilingualism.

Optionality and learnability

As stated, if a grammar must either be +Agreement or -Agreement, then a single grammar cannot allow both "I want" and "me want". Under the TB approach, the child is never required to convert an optional rule into an obligatory rule.[12] Instead one grammar is abandoned. This is a step forward because it solves a traditional puzzle: it is very difficult to imagine the evidence that would force conversion of an optional rule into an obligatory rule.[13] If Agreement is optional, then hearing an example like "he walks" cannot establish that it is obligatory.[14]

The link to social registers

"Pro-drop" languages allow null subjects ("goes" instead of "he goes") and they are commonly differ-

[12] See Wexler and Culicover (1980) for early discussion of this question.
[13] See Berwick (1985) and the learnability literature.
[14] This observation is pertinent to those dialects, such as African-American English, in which Agreement does not always occur. It is a well-known phenomenon in speech pathology.

entiated from languages which have obligatory subjects. And yet in English one can, in an informal social register, delete matrix subjects with certain verbs (''seems like a good idea''/''looks good to me'').[15] The missing subject is either a special rule, called ''Diary Drop'' (Haegeman, 1990), or it is the marginal presence of ''pro-drop'' in a non-pro-drop language. In either case, it is a radical departure from the usual obligatory subject requirement.[16] What is of interest is (i) that the choice of grammar can be linked to social register, and (ii) that the social register feature varies independently of the grammatical structure. Subject deletion is not necessarily informal in Romance languages.

One is led to this hypothesis: a shift in grammar signals a shift in social register. It is precisely because a principle from another grammar system (or a default economical system) is used that a shift in social register is communicated. For instance, we can sound biblical or Shakespearean by using features of Old English that are Germanic in origin. Relics of a productive rule of WH-movement inside PP's produces forms like:

(6) whereafter
    wherefrom
    whereunder
    wherewith

This is not completely general:

(7) *wherearound
    *whereamong
    *wherethrough

If we say ''whereafter'' it has a formal, almost legalistic, tone in modern American society, while it may have been without that overtone in earlier periods of the language. There is no prepositional pre-posing rule in modern English, probably because there is no ''prepositional complementizer'' in modern English, while older forms of the language allowed the projection of an additional structural layer, or perhaps an even more complex mechanism. It seems here that what makes one social register distinctive is that it exhibits basic operations that belong to a different grammar.

We will extend this approach to domains within adult grammar in which we argue that grammatically incompatible forms coexist only because the speaker is ''bilingual''. For instance, as we argue below, an English speaker can use Germanic V-2 structures as a mode of social emphasis.

Theoretical sketch

We provide here a perspective on the relations between principles of economy, a Default Grammar, and a particular grammar.[17] This is then the formal source of one form of bilingualism:

(a) Universal Grammar defines a set of default representation which all speakers possess. We call this: Minimal Default Grammar (MDG).
(b) The set of MDG structures reflects principles of economy. That is, they project fewer nodes than elaborated particular grammars.
(c) The particular grammars and the MDG grammar may or may not be incompatible.
(d) Different grammars can be localized:
    (i) in lexical classes
    (ii) by speech register

The notion of MDG in (b) captures the universal structures which contain no language-particular information. For instance, the Determiner Phrases vary from language to language in how much Agreement they contain, while (possibly) NPs below DPs are completely universal. Similarly, the notion of incompatibility in (c) follows directly if, for instance, Agreement is obligatory in a particular language but not present in the MDG representation.[18] If a grammar lacks Agreement, then it is a direct reflection of MDG.

Lexically restricted V-2 in English

The first form of bilingualism we consider is linked to the lexicon and not linked to principles of economy. Suppose I say the following seemingly anomalous sentence, which some readers will recognize, not as a fixed idiom, but a kind of ''idiomatic style of locution'':

[15] Chomsky (pc) has suggested that pro-drop is linked to speech register.

[16] Observations of this kind have motivated the idea that constraints are universal in Optimality Theory. Default Grammars bear a similarity to Optimality Theory in this respect. However the notion that bilingualism is universal does not fit the notion of ranking which is used to differentiate languages in Optimality Theory. In other words, under OT, as in the Minimalist program, there is no reason, given only one grammar, that all traces of a different grammar would not be driven out.

[17] Vainikka (1990) and Lebeaux (1990) initially introduced the notion of a default as an important aspect of acquisition. See their work for other relevant formalization and observations. See De Villiers and Roeper (1992) for use of the notion of Default case, and more recently Schütze and Wexler (1996).

[18] Roeper and Rohrbacher (1994) argue for precisely this view, based on Speas (1994) who argues for the optionality of Agreement. See also Chomsky (1995) who reduces Agreement to a feature on the Tense Phrase. And see Schütze and Wexler (1996) who extend the argument for the optionality of Agreement.

(8) A single salad does not a dinner make.

This form is generalizable:

(9) One captured fish does not a fisherman make.

Clearly we have a sort of an idiom with some *lexical openings* into which we can put virtually anything (salad, dinner, fisherman). Is there any significance to this idiom that is unlike any other idiom?

The special feature of this idiom is that it uses an operation which is at the heart of many Germanic languages, but not English. We will begin with an informal version of the rule and progressively refine it:

(10) Put the main verb in final position

The verb final structure is also associated with a special movement rule, known as Verb-second:

(11) Move the verb directly into second position, i.e. the complementizer position.

Such movement of the main verb was present in Shakespearean times and continues to exist as an idiom in modern English.

(12) Say you so?

The rule allows movement of the main verb beyond a Negative Phrase as well, and this appears in other current idioms:

(13) It matters not what you do

(13) has exactly the same meaning, but not the same impact as the non-idiom form (14):

(14) It does not matter what you do.

We must ask why we should have a second form, with the same meaning, that appears to travel back centuries in the history of the language to a point where a different verb-final "deep structure" is present.

Before we proceed, we must observe that each of these expressions has distinct limitations. The nouns can be freely exchanged but the verbs are quite limited:

(15) (a) A dessert alone does not a meal make.
(b) Think you so?
(c) ??Believe you so?
(d) *A tiny orange does not someone peel.

Although (d) has virtually the same structure, it no longer feels like an idiom. So we have two features, Verb-final structure and V-2 movement, which come from Germanic languages and define a family of idiomatic structures in English. Are they just complex lexical items? Are they add-on rules to the existing rules of English? In principle they cannot be added on to English because they are in a sense "at odds with the deep structure of the language". English is SVO and German is SOV. Thus we might argue for a deep structure bilingualism principle:

(16) (a) Any rule compatible with one deep structure can belong within one grammar.
(b) Any rule which presupposes a different deep structure belongs to a different grammar.

Although current theories lead to a more intricate formulation, as we discuss shortly, this remains a reasonable hypothesis.[19]

The representation of V-2 in the adult grammar is sharply limited to a specific set of verbs. Next we turn to the acquisition question: how does the English child decide to adhere to a highly limited rule, while the German child decides to make a fully productive rule?

Acquisition

Evidence for V-2 in English extends beyond a few main verbs. The verbs "be" or "have" operate as main verbs which undergo V-2. They are so frequent that one must ask why they do not trigger V-2 as a general property of English. Given the child's gradual exposure to the language, this is a logically significant possibility. We find that both "be" and "have" invert:

(17) (a) is he here
(b) have you a dollar[20]

In sheer frequency terms, the child hears a significant portion of V-2 expressions (like "what is that?").[21] In

[19] A current theory by Kayne (1994) suggests that even this distinction is rule-governed: all languages are SVO but some overtly move the object over the verb in order to receive case in a higher "functional" category and others do so covertly (invisible movement occurs for certain elements (see Chomsky, 1995). Now the distinction is narrower: one rule applies in German but not in English, except in idioms.

This new version of the Universal Base Hypothesis suggests that languages are closer to one another than they first seem and they make it natural that a set of idioms in one language might mimic the grammar of another language. One language allows a subset of lexically defined items to undergo an extra rule. This conception makes the notion of a distinct language as an object more obscure from a formal perspective. It seems that all possible languages projected by UG are generable by rule from each other. In the extreme form then, every language could just select options, word by word, from UG. The proportions would vary drastically: English has a few V-final structures and German has thousands.

[20] This form is becoming fairly rare in modern American English, but less so in British English.

[21] See Takahashi (1989) and Stromswold (1995) for arguments that inversion must be present in these cases. Note that demon-

order not to mis-set the V-2 parameter, the child must retain a lexical connection. Without a lexical connection, the child is exposed to two grammars, V-2 (''what is that'') and non-V-2 (''what did he say not'' ''*what said he''). One would therefore expect the child to be paralyzed, unable to choose, faced with an unlearnable grammar. Instead of paralysis, TB enables the child to choose both.

In addition, the entire class of speaking verbs allows V-2 in quotation environments:

(18) (a) ''Nothing'' said John
(b) ''Go'' shrieked the witch

The verbs ''say'' and ''shriek'' have moved beyond the subject here. Children's stories, often repeated, are full of quotation inversion – see Collins (1997) for discussion – and it is ungrammatical to say:

(19) *''Nothing'' did John say.

How does the child determine that it is just in this domain that V-2 is allowed and must not be generalized? The German child by contrast decides that V-2 is general.

There is subtle and brief evidence that children (a) attempt to treat ''have'' and ''be'' like other main verbs that do not invert, and at a different point (b) attempt to expand the set of V-2 verbs which do invert. Each of the opposite rules generalizes slightly beyond the specific lexical types given. For a stage that may be as brief as a week, children sometimes utilize do-insertion to prevent the inversion of ''be'':[22]

(20) do it be colored
you don't be quiet.
Allison didn't be mad
this didn't be colored
did there be some
does it be on every day ...
does the fire be on every day
do clowns be a boy or a girl

English cannot be simultaneously V-2 and non-V-2. The conflict can be managed only by linking V-2 instances to the lexicon.

The lexical link does not mean that the child proceeds on a purely word-by-word basis. Children, like adults, must allow quotation inversion to include the whole class of verbs of speaking (''mutter, shriek, announce'', etc). There is a small amount of evidence[23] that children will use lexical class as the basis of a V-2 generalization. For a few weeks one child consistently uttered sentences of the form in (21):

(21) what means that [instead of ''what does it mean'']
what calls that [instead of ''what is it called'']

The verbs ''call'' and ''mean'' both fit roughly within the class of *equative* verbs (''be, equal, constitute''). In sum, from an early moment, children circumscribe the V-2 option in lexical terms, although they receive substantial input which is compatible with it and therefore one might expect the child to generalize to a full V-2 operation.

The evidence for ''undergeneralization'' in children is widespread. They do not take every new word which has a distinct rule and extend the rule to all other words. Thus the grammar is lexically conservative. This leads to the following picture:

(22) Hypothesis: Children establish vocabulary sets which are independently derived from principles of UG. Each subvocabulary set follows its own rules.
Consequence: Two lexical sets constitute two grammars.

This is a strong view of inherent bilingualism in all speakers. Without such a possibility, English could not maintain distinctive subvocabularies of Anglo-Saxon, Latin and Greek origin.[24] We have now defined one form of Theoretical Bilingualism which is localized in lexical classes and which reflects the process of historical change. English evolved from a

stratives cannot function as predicates: ''*a fish is that''. Therefore ''what is that'' must come from ''that is WH-something''.

[22] See Roeper (1993) and Davis (1987) and references cited therein for sources. Moreover, adults using American English today are progressively avoiding inversion with ''have'', preferring (i):
i) Do you have a dollar?
We are in the midst of a form of language change with respect to the verb ''have'', which notably has the social register characteristics under discussion. Every speaker, I think, would say '' do you have a dollar'' feels more informal than ''have you a dollar''. The fact that the change comes slowly reflects the central thesis of this paper that bilingualism is present in the adult language: the adult has both representations of ''have'' as undergoing V-2 and not undergoing V-2.
It is demonstrably not the case that children allow other auxiliaries to be treated as main verbs. If they did, then we would expect main verb usages to appear, which are common in other languages where modals are main verbs. However I have never heard of an English-speaking child saying (i), although (ii) is common in German:
i. *I can everything
ii. ich kann alles (''I can everything'')
Therefore the application of V-2 to main verb ''have'' is strictly limited lexically.

[23] This comes from my personal diary evidence from Tim Roeper.

[24] For instance, see Randall (1980) for a discussion of affixation. She shows that speakers know that ''civility'' is possible but ''*evility'' is not since the latter is Anglo-Saxon and not Latinate. However the Anglo-Saxon affix-''ness'' can appear with both forms: ''civilness'' and ''evilness'' How did ''-ness'' lose its Anglo-Saxon moorings and become productive for all nouns?

V-2 language and retains a subvocabulary which continues to adhere to that grammar.

Many mysteries remain about how and why languages change. The potential for universal bilingualism explains in part how such changes can be gradual. The largest historical mystery is how one lexical class becomes productive and the other remains unproductive. The same mystery arises in acquisition: at what point does one lexical class, linked to one grammar, become productive and dominate the language? At some point the grammar becomes more abstract. It restates a rule that is marked V-Latin to simply V, but we do not yet have the formal insight needed to state this shift correctly.

We turn now to a re-examination of this same question from the perspective of language interference. Our discussion will engage more modern versions of V-2.

Language interference

Is there an abstract answer to the question: how can grammars interfere with each other? Code-switching and lexical borrowing constitute evidence of where grammars can connect and interpenetrate. But we do not know, offhand, if such connections are accidental or conform to principle. Speakers sense subtler influences as well. It is a very interesting theoretical question: where are dialects open to influence and how is this influence manifested? Phonologically, it is clear in various accents that certain distinctions may be lost. While phonology may help to keep grammars distinct, interpenetration is certainly evident. In syntax, the influence may be less manifest. Consider this hypothesis about interpenetration:

(23) Grammars may not be distinguished by bilingual speakers if they differ only in the overt/covert status of an operation.

We shall argue, however, that perhaps no rules have such a minimal distinction: all movement is accompanied by some semantic distinction (which may force movement in order to satisfy checking).

Let us consider one famous case. Chomsky (1995) proposes that the V-2/non-V-2 difference involves only Phonetic Form: V-2 is overt in some languages (German) but occurs covertly in others. Verb-raising is obligatory in all languages in order to check off tense features. Nevertheless, V-2 is not identical in English and German for two reasons: (i) the operation occurs overtly in German, but not in English; and (ii) movement appears to go further to a CP node in German which in turn allows inversion structures not available in English (*toast eats John).

The first distinction is the famous distinction motivating the work of Pollock (1989) in which the fact that verbs move over adverbs in French, but not in English, is explained by the absence of movement in English. Chomsky (1995) argues that the movement still occurs, but at a covert level because all verbs must be linked to tense features for interpretive purposes.

This syntactic explanation, however, does not capture all of the grammar differences. We claim that an important, though subtle, semantic difference exists between overt and covert raising, which has not been integrated into syntax before. English, notoriously, has "no present tense" which is an informal way of stating the surprising fact that the grammatical Present in English cannot refer to the actual present, but must refer to the generic:

(24) John sings

does not entail the present:

(25) John is singing

It asserts only that John has the ability to sing in general with no commitment about the present. In German, however, the present, which overtly raises in V-2, is ambiguous between the meanings of (a) and (b):

(26) Hans singt = John sings or John is singing

It cannot be a coincidence that just in the language where there are "weak" features, we find an absence of temporal anchoring, or finiteness. It suggests that raising Checks off two features: Tense and Finiteness. Where raising does not occur overtly, then finiteness is not fixed.[25] This perspective can provide a deeper reason for the weak/strong distinction and the existence of overt/covert movement. The deeper argument is that overt movement of all kinds is a device to achieve the property known as *visibility* which is associated with definite reference for Noun Phrases. We now argue that visible movement gives definite reference, via temporal anchoring, to Verb Phrases.

[25] Meisel (1994) represents Tense as distinct from Finiteness, locating Finiteness in C, following Holmberg (1986), and Håkannson (1998) argues that children fail to represent Finiteness as opposed to Tense. Moreover, Herschensohn (1998) provides evidence that in L2 raising is acquired in a lexically-linked way with specific verbs shifting to Raising. She provides no discussion of the Finiteness factor.

Wexler (1998) argues for a "unique checking" limit within a grammar that allows a child to check either Agreement or Tense, which in turn can lead to either nominative or accusative. His approach would effectively build two grammars in one in order to maintain a single grammar theory. While one might construe these as notational variants, one would look for a distinguishing factor under the TB approach, rather than the assumption that variation is arbitrary.

If two grammars are involved, then we can predict that the same distinction will arise in the exceptional V-2 lexical class of speaking verbs. Though subtle, we believe that the prediction is upheld:

(27) (a) Here's what happened. Bill comes in the room with a new toy.
"Awesome" says John over and over.

The inverted structure refers to a single event. Were one not to invert, then the dialogue becomes strange:

(27) (b) Bill comes in the room with a new toy.
John says "awesome" over and over.

In the inverted form (27a), finiteness is implied and only one event has occurred, perhaps in the narrative present where a story is being retold. In (27b) the uninverted verb carries the generic reading and means that John characteristically says "awesome". Therefore we find that the fine structure of the language is obeyed in these contexts.[26] The Germanic tense-anchoring linked to V-2 is found in the English subvocabulary that permits V-2.

R. Schafer (pc) has noted a similar effect with auxiliary raising over an adverb:

(28) (a) The children already have gone to see Robin Hood
(b) The children have already gone to see Robin Hood

Most speakers, when asked, will take (28b), where "have" has raised above the adverb "already", to mean that the children are not here right now because they are at the movies, while (28a) means that they have seen the movie sometime in the past. Thus the movement of the auxiliary "have" anchors the past tense, just like verb movement anchors the present. Therefore the Finiteness feature may remain an ingredient in residual V-2 as well.[27]

Nevertheless, the Finiteness or Temporal Anchoring feature appears to be one that can affect other grammars, that is, interpenetration occurs. It is often observed that non-native speakers of English have difficulty in (i) overuse of the progressive, or (ii) misuse of the present to indicate a current activity. Thus one might hear the dialogue: "where is John?" with the answer "he sings" when the intended meaning is "he is singing". Thus the L2 speaker has either incorrectly imposed a Finiteness feature on the unraised English verb, or in fact raised the verb to acquire Finiteness when it does not raise in English. How can the L2 speaker allow this to occur? The fact that raising is invisible in many sentences means that the German speaker could raise the verb in "John sings" while the English speaker does not and there would be no overt evidence to the contrary. This is then an example of how we may find grammar interpenetration just at the point where the overt/covert distinction applies.

In what follows we will define a second origin for universal bilingualism in terms of economy.

Minimal default grammar and economy

One feature of economy in Chomsky (1995) is *economy of representation*:

(29) Project minimal amounts of structure.

The claim in (29) is a programmatic suggestion that must be analyzed in terms of language diversity.[28] Whatever is a universal requirement of all languages cannot be omitted. Therefore each claim of minimalism must be defended. For instance, if Determiner Phrases are universally present above Noun Phrases, then they should not be omitted, but if languages allow NP to occur by itself, then (29) predicts that it should be the first hypothesis.[29]

First Vainikka (1990) then Lebeaux (1990), and Roeper and deVilliers (1992) have pursued the idea that there are default structures to which children have access. These two strands lead to a natural combined hypothesis:

(30) Default structures are defined as economical structures (Minimal Default Grammar (MDG))

The characteristic feature of defaults is that they can be projected with no direct input. They are generated directly by Universal Grammar.[30] Therefore, as we argued above, sentences of the form "me want" arise among a number of English-speaking children when they recognize "me" as the default case form although adults never say "me want". We have argued that a more economical representation, no AGR feature, leads to this possibility. Since children simultaneously use both "I want" and "me want", the Minimal Default Grammar introduces another form of bilingualism.

[26] Tamanji (1998) extends this view in a number of ways, in particular to movement in an African Grasslands language, Bafut, where verb-movement exists which is not movement to Tense.

[27] An anonymous reviewer points out that weak verbs optionally raise in French. Our argument suggests that one should seek subtle semantic effects of such movement.

[28] The economy of representation approach is pursued in work by Roeper (1996) and Roeper and Rohrbacher (1994).

[29] See De Villiers and Roeper (1995) for discussion.

[30] Therefore they have properties like those found in creole languages discussed by Bickerton (1981).

Hypothesis (30) leads to the view that we can use properties of child grammars to define features of UG. In this instance, it suggests that we define the notion of economy so that it predicts the default structures which have been observed. For instance, resumptive pronouns are found in many dimensions of child language. There are many examples of resumptives in child language (see Labelle, 1990, and Perez-Leroux, 1995):

(31) here's a little kid that he talks
I hurt my finger that Thomas stepped on it
you are a tree and I'm a kid that I climb up on you
Smokey is an engine that he pulls a train
twenty$_i$ numbers that we counted them$_i$ [31]
(from D. Finer, quoted in Perez-Leroux)

The presence of such structures in child language then requires that we state a form of economy which says, roughly:

(32) (a) Pronominal indexing is more economical than
(b) movement operations

Therefore the grammar prefers (32a) to (32b), but one must now seek a formal representation that leads to the same conclusion. We will not pursue this modification of economy in detail at this point, but the approach should be clear.

Tense chains and economy of representation

We turn now to a notion of economical representation which derives from acquisition and second-language phenomena. However, it requires an economical representation not of structure itself, but economy in the application of a Principle, c-command.

A current issue in modern grammar is the explanation of the phenomenon of *do*-insertion. Why and where does it exist? Chomsky (1989) has argued that *do*-insertion is a Last Resort operation when movement of the main verb to Tense fails. We will not provide a full analysis of this phenomenon, because it is quite complex, but rather explore one prediction and one form of economy of representation to which it is linked.

In recent work with Bart Hollebrandse (Hollebrandse and Roeper, 1996), we have argued that *do*-insertion should be analyzed as what is regarded as a strong affix. Once again, grammars divide into those with a weak affix system, like English, and those with a strong affix system, like Italian. The strong affix can appear independently in an Inflection node. The weak affix, by hypothesis, is linked to the verb in the lexicon and is inserted under the V-node together with a verb. Then it moves higher to the Tense node position. We argued above that this movement may be analyzed as involving the absence of a Finiteness feature for the weak form.

We argue, however, that *do*-insertion is just the spellout form of a strong affix. In other words, the form "did" is just the way we pronounce -ed by itself (following a suggestion by H. Lasnik (pc)). Under this hypothesis, however, English contains both strong independent affixes linked to "do" and weak affixes which are generated as a part of the verb. Therefore, once again, we have a hidden form of Theoretical Bilingualism.

English provides the child with mixed information in this respect. We find that the strong affix is used in questions and negation, but not in declaratives (33e):

(33) (a) did he talk
(b) he did not talk
(c) *talked he
(d) *he talked not
(e) he talked

Hollebrandse and Roeper argue that the *do*-insertion form is in fact preferable.[32] In effect, then, it is a First Resort phenomenon rather than Last Resort, because it obeys principles of economy, as we shall show. From an intuitive perspective, the argument is this: the tense marker in "talked" is buried in the verb, while the tense marker in "did talk" is explicit.[33]

In formal terms this idea can be expressed in terms of a refined principle of economy applied to trees. We assume, following Guéron and Hoekstra (1988), that Tense and the verb are linked by a tense chain which requires that the higher tense marker dominate or more precisely c-command the lower verb. The chain is visible in speech errors, common among L2 speakers, who link both weak and strong in forms like "did he left".

Now we argue for a narrower notion of

[31] Note that the view that this is purely a processing effect would not explain sensitivity to quantification.

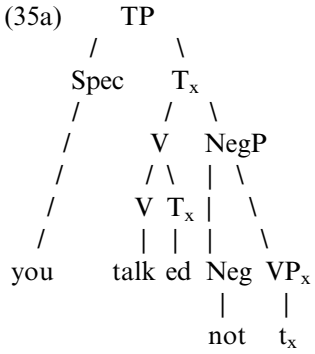
[32] See also Cavar and Wilder (1996) for similar arguments applied to Serbo-Croat.

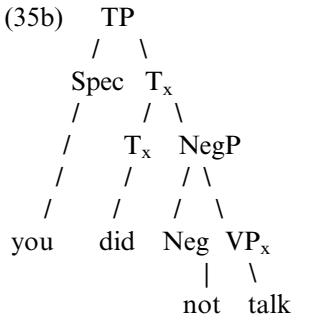
[33] See Ravem (1978):
Subject: Reidun (3;9 years old); native speaker of Norwegian.
Examples: I did bit it
Cause I did want to.
We did saw that in the shop.
I did shut that careful.
My mummy did make lunch for them.
Whos did drive to Colchester? (subject-WH monoclausal Questions)
Ravem reported that "did" is not an emphatic form in these utterances. The error is common among L2 speakers.

c-command as the default form, in which the morphological affix -ed (pronounced as "did") directly c-commands the verb (in an x-chain). Lasnik (pc) has argued that "did" is the spellout of a past tense feature. Therefore we have in effect *feature-command*:

(34) C-command should be morphologically direct.

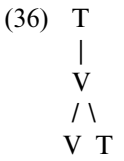
This can be illustrated in tree-form. In (35a) the T (Tense node) dominates a V which dominates another T, while in (35b) T dominates T directly.

(35a)


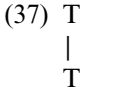
(35b)


←covert————————talk

In effect, the grammar must look down from the T-node into a V node to find another T element :

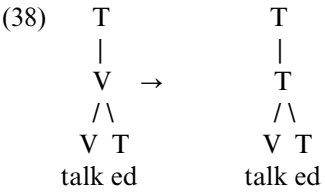
(36)


As opposed to a direct link (37):

(37)


Where the direct link is present, the morpheme -ed directly c-commands the main verb node to which it is linked (x-chain).

How does the grammar "look down" in (35a)? Chomsky (1995) suggests that a higher node can "see" the nodes below it and therefore no difficulty is present.[34] Hollebrandse and Roeper (1996) argue that the distance downwards to the crucial Tense -ed feature makes an economy difference. Therefore if the child hears both "talked" and "did talk" she can immediately recognize that the latter creates a more economical chain because it involves a *shorter downward distance* to locate the tense feature under the T node and conversely a direct c-command relation over the lower verb. They suggest that for "talked" one must relabel the V to a T-node in order to allow the feature to percolate to the higher T-node:

(38)


Evidence that the "look down" mechanism is real is reflected in the fact (K. Johnson, pc.) that certain verbs require immediate domination in their subcategorization:

(39) (a) I wondered who I saw a picture of
(b) *I wondered a picture of whom I saw

In (39b) the WH-feature is not directly dominated by "wonder".

There are, in fact, a variety of technical options for refining the feature-checking mechanism. Our goal here is simply to argue that "did talk" is simpler than "talked" for purposes of feature-checking.

If we are correct in arguing that a form of economy is present in do-insertion, then we predict that children can spontaneously project *do*-insertion forms. Exactly this occurs in both English and Swiss German (see also Penner, 1994). Thus we find (without any emphatic stress) (40a) and tense-doubled forms (40b):

(40) (a) I do have juice in my cup
I do taste them
I did wear Bea's helmet
I did paint yellow right here. I did put the brush in.
I did paint it
what did take this off
do it be colored
does it be on every day
did there be some
A doggie did walk with Dorothy and the Doggie did hurt itself

[34] See Roeper and Perez (1997) for further discussion of how non-c-command relations interact with Pied-Piping in early grammars.

(40) (b) I did broke it[35]
I did fell when I got blood
I did fixed it
Jenny did left with Daddy
I did rode my bike

The double-tensed forms appear is found not only among children but very frequently among L2 speakers.

"Do" in German and Dutch acquisition

This form also appears briefly in Dutch and German child language where it is common among dialects and may occur in parent-child language.

(41) ik doe ook verven
"I do also paint"
ik does grapjes makken
"I do grapes make"
hij doet taperecorder draaien
"he does taperecorder turn"
wat doet 'ie bukken
"what does he stoop" (CHILDES) (from van Kampen, 1996)
wat doe jij zeggen
"what do you say"
dat doe ik spelen
"that do I play"

We now make an additional prediction, namely, that the reverse never occurs. There are no reported examples of children who say:

(42) *John talked not
*Bill sang not
*what bought John

There are exceptions to this claim which are precisely the V-2 structures noted above in lexically restricted classes "what means that".

If we combine our two examples we make a further prediction:

(43) Children make anti-economical overgeneralizations only in lexically defined ways.

[35] Pinker (1984) notes that these tense-copying environments are more frequent, but not exclusively, associated with strong verbs. The fact that strong verbs are involved means that the actual system of tense-agreement linked to lexical lookup may be slightly more complex in the adult grammar and therefore have an impact on the child grammar. The fact that the phenomenon also occurs with non-strong verbs means that our analysis still appears to be on the right track. The alternative is to argue that the notion of past is incorporated lexically in a way that makes it inaccessible and irrelevant to tense-agreement. It is not, for instance, the case that we do tense-agreement with adverbs such that "was+today → yesterday". Instead we mark tense on both the verb and the adverb (was, yesterday) independently.

Conversely, only forms defined within MDG will overgeneralize beyond lexical classes.[36] Now we can apply the same argument to some of the V-2 examples we have seen. In essence we argue that when the child is exposed to both forms:

(44) (a) what had you
(b) what did you have

the child will recognize (44b) as being more economical than (44a) because the tense chain obeys c-command directly. It is now natural to argue that V-2 will arise in lexically limited ways for both L1 and L2 learners (as Herschensohn, 1998, argues), because V-2, failing to be economical with respect to c-command, is inherently marked. This hypothesis (43) is one, traditional, view of exceptionality, locating it in the lexicon. In the next section, we will propose a stronger principle to explain why two rules may fail to collapse.

Incompatible economies

What is the connection between the arguments we have presented and historical linguistics? In a sense, the question of change over time is the logically subsequent question to the question of how to represent grammars in conflict. Why do some parts of the language yield to change in the direction of uniformity and others remain immune to change?

Kroch and Taylor (1997) summarize a series of papers which detail the gradual shift from V-2 to lack of V-2 in the history of Germanic. A huge roster of factors seems relevant, far beyond what we can consider. They show an apparent (and perhaps ultimately real) gradualism in the shift away from V2 with respect to pronouns, PPs, and topicalized NPs. (e.g. "the hat I saw"/"the hat saw I")

We shall not probe those mysteries, but rather limit ourselves to seeking to represent and explain one domain where "two grammars" resist the pressure to collapse into one. Why does the quotation remain one domain which resists a shift to V2? What guarantees its stability?

Here, again, is the essence of the situation. Quotation optionally allows inversion:

(45) (a) "Nothing" John said
(b) "Nothing" said John

but does not allow just the auxiliary to invert:

[36] Our discussion has not differentiated movement to IP and movement to CP, which have been classically regarded as a decisive difference between English and Germanic. Recent analyses have in fact suggestion that Germanic languages also involve movement to IP (Zwart, 1993). The core arguments here go through if we further differentiate landing sites for questions as opposed to declaratives (IP and CP).

(46) *"nothing" did John say

Why is auxiliary inversion insufficient? In contrast, question formation and locative inversion with polarity items obligatorily requires inversion, but only of the auxiliary ("residual V2"):

(47) (a) what did John say
(b) *what said John
(c) No one did John see
(d) *No one saw John

Where non-polarity items are involved, we get both forms:

(48) (a) into the house John went
(b) into the house went John

It is these latter cases which seem to be subject to gradual change in the data of Kroch and Taylor (1997).[37]

Why is quotation immune to change? If we follow the reasoning of Yang (1999) who argues on learnability grounds that children seek "local maxima" allowing grammars to remain in conflict if there is sufficient justification for each case, then we may be able to appeal to the idea that each grammar has achieved an independent form of economy.

We will sketch an analysis of each form.[38] First, as we argued above, the movement of the auxiliary, but not the whole verb, preserves one form of economy:

(49) Direct feature-command is economical

Therefore the tense chain is economically preserved if only an auxiliary "do" is projected

(50) what did_i John t_i say_i

This chain also involves a checking relation with a quantificational feature in the polarity item ("no one") or WH-word. Therefore inversion is obligatory in cases such as:

(51) No one did I see

Now we must ask why this should not be sufficient for quotation.

The core reason, intuitively, is that quotation can be fixed in the Here and Now only when the verb raises. This predicts that it is impossible to have the progressive as a source of temporal anchoring for quotation. This is correct:

(52) *"yes" is John saying.

Now we will represent this claim in a more formal discussion.[39] Temporal anchoring is a form of specificity of the same kind that is indicated for NPs or DPs. Following Collins (1997) we imagine that there is a Quotation Operator in CP which requires independent checking.

(53) We suggest that:
there is a specificity feature on the quotation, like a DP,
which must be checked by a [+Quotation] Operator feature on the verb[40]

The specificity feature is linked to a Quotation Operator that is linked to, but not the same as the tense feature. We have argued above that failure to move the verb overtly will fail to achieve temporal anchoring, which is now translated into checking a specificity feature. Movement of the verb overtly instead of covertly achieves *Local Economy*, because the formal features are in a spec-head relation rather than depending upon a covert chain into the VP.

Can we find this effect of verbs elsewhere? Note the specificity effect of a full verb in ellipsis:

(54) (a) John pushed his car and Bill pushed too → specific object
(Bill pushed John's car)
(b) John pushed his car and Bill did too → sloppy reading
(Bill pushed Bill's car)

In (a) Bill pushes John's car, while in (b) we get a sloppy reading and Bill could push his own car.

Local economy is maintained if the specificity requirement is fixed overtly by the moved verb?[41] Thus we have:

[37] Müller (1998) makes the plausible and interesting claim that transfer occurs at points of ambiguity. The question which then arises is how to define ambiguity. If "be" raises in English, then is it evidence for V2 or residual V2? The answer depends on whether "be" itself is seen as a main verb or an auxiliary.

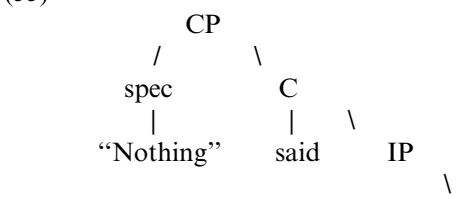
[38] The pertinent argumentation is far more intricate. We refer the reader to Collins (1997) whose analysis we follow with the addition of the specificity concept to which we turn directly.

[39] A similar distinction is subtly evident in the presence of both direct and indirect question formation in English. It happens that people will say either (i) or (ii) with or without inversion, although (i) is judged to more grammatical:
i. John wondered which song he should sing
ii. John wondered which song should he sing
In (i) the assumption is that there is a fixed array of songs from which he should choose. In (ii) the implication is that John is seeking to make a choice from an unfixed potentially infinite array.

[40] See Collins for an explanation of the Quotation Operator and uninverted cases ("Nothing" Bill said) in terms of Object Shift.

[41] The temporal anchoring property provides an explanation to what Collins says is a stipulation in his theory:
The EPP feature of T may enter into a checking relation with the quotative operator only if V[Quote] adjoins to T. The intuition behind this stipulation is that T must be supported by the actual quotative verb in order to check the D[quote] feature of the quotative Operator.
In effect, then, this is a more technical formulation of our earlier proposal that verb-raising is linked to temporal anchoring, but now applied to the quoted material itself.

(55)



If quotative V2 is justified by specificity features which must be checked by movement, then why not assimilate "residual V2" to full verb inversion: eliminate do-insertion. Put differently, why would history not go backwards? The answer lies in the fact that the emergence of residual V2 allowed an economical tense feature chain. The child prefers to keep two grammars if this principle is contravened:

(56) Two grammars will not assimilate if the elimination of a more economical representation in either grammar is required.

This is like the suggestion by Yang (1999) that *local maxima* exist which are incompatible, but since each receives sufficient support, they remain in a "steady state". [42]

This line of reasoning will explain why a language will tolerate incompatible domains in the grammar, but not why language would change at all. The answer may lie with how languages shift at a deeper level not captured by this kind of formalism. For instance, the shift from a tense-dominant to an aspect-dominant language is not easily expressible in this system.[43]

Mysteries remain

To my mind, the foregoing discussion marks a viable form of progress both in the application of linguistic theory to problems of bilingualism, and, in turn, in making linguistic theory responsive to the large range of provocative data that is currently emerging from work in first-language acquisition, bilingualism, second-language acquisition, and communication disorders.

Nevertheless, we must emphasize that fundamental questions remain unanswered:

(57) Non-economy:
Why do non-economical forms exist at all? In current theory there is no reason for the presence of V-2 at all, since feature-checking at LF supposedly can achieve the same result. We begin to decompose this picture via our proposal that overt movement is required for tense anchoring.

(58) Acquisition:
We cannot state exactly why the Germanic child does not arrive at the same conclusions as the English child, i.e. the same language, given evidence that *do*-insertion and its economic advantage are present in those languages at certain points in the acquisition process, i.e. both German children and Dutch children pass through a stage where they use do-insertion.

(59) Productivity:
Finally we are left with one of the deepest mysteries in linguistics: when does a rule become productive, when does it lose productivity, what keeps a rule bound to a lexical island? These questions are linked to the question of historical change. They remain deeply puzzling. Why does do-insertion suddenly emerge in Middle English and why does it emerge and then leave child Dutch and why is it briefly over-productive in English?

Are there deeply formal answers to these questions, or should we look at an interface between social register and grammar? Is it some social nuance in language that suddenly gives a certain rule prominence?

A speculation

Why should we ever move the full verb when presence of a c-commanding tense morpheme (or even an invisible feature) is sufficient? We have argued that V-2, unlike English, checks a Finiteness feature, but one must still ask: why not capture this feature with a minimal verb, as in the English progressive?

The explanation for V-2 is a prominent puzzle that has been addressed in the Minimalist Program by many scholars.[44] One possible answer to this question lies in the notion of economy linked to modularity. Consider this hypothesis:

(60) Economy exists independently in different modules.

Suppose further:

[42] Yang approaches these questions partly in terms of frequency which we continue to avoid.

[43] The temporal anchoring accomplished by moving the main verb is now accomplished by the verb combinationin "is running". This seems indirect and almost misleading because progressivity seems incompatible with stativity. The expression "the birch tree is standing in the corner of the yard" seems to imply an ongoing activity, rather than a state. Clearly there is a deeper system of compensatory change taking place that we do not yet grasp.

[44] Discussed in Chomsky's Fall 1995 class.

(61) No LF operations occur inside words, therefore morphological economy requires an adjacent, linear array that matches the UG specified order of interpretation.

Strict morphological ordering of verbal morphemes is typically reflected in heavily morphological languages (see Baker, 1988). Ordering within morphology is very strict in the derivational realm. Consider a simple case: *destructiveness* versus **destructnessive*. Baker (1988) has argued that similar constraints hold for syntactic morphemes.[45] In fact, the debate over how Agreement, Tense and Aspect are ordered partly involves their morphological order. If we argue that the morphological principles require Verb+Tense to be interpreted before Verb+AGR, then the interpretation is matched by the morphological sequence: in German Tense is inside Agreement (see Meisel, 1994 for extensive discussion):

(62) sagtest = sag + Tense + AGR
te – st

Using do-spellout to create a tense chain, obscures the relation of tense to other verbal morphemes. The order of morphemes and verb is preserved directly if the whole verbal complex is fixed in an adjacent array, via verb raising, but it would not be preserved if the tense morpheme is detached. We could then reconstruct a chain TP ... VP with no ordering; one could construe that Tense+verb, or verb+Tense, while with the moved verb we have a fixed order: verb+Tense, or if AGR is a separate node: verb-Tense-Agr.[46]

Achievement of a strict order that suits interpretation within morphology is accomplished by overt movement where the hierarchical order is syntactically fixed. Therefore morphological economy invites V-2. This is a more refined view of what is known as Holmberg's generalization, that rich morphology correlates with V-2. We argue that it is the internal structure of morphology which leads to this consequence. This is merely a suggestion which does not confront many intricate aspects of the morphology/syntax interface.

Now we have a paradox: raising an auxiliary gives us economy of feature-command. And raising the main verb gives a direct reflection of LF in the AGR and Tense sequence. Each kind of economy destroys the other.

Speech registers

Why do languages have pockets of TB? This would seem to be highly inefficient from a formal point of view. The answer, as we hinted above, may lie outside formal linguistics.

What makes a *social register* distinctive? What conveys to people the sense that a different level of communication is involved if, among bilingual speakers, one or the other language is chosen? These are deep questions which go beyond linguistics and my realm of expertise.

If we follow the logic of this essay, then a straightforward hypothesis arises, namely that a speech register has a formal dimension:

(63) Formal or informal speech registers are recognizeable as a choice of a different application of principles within UG.

If the normal register does not allow preposing inside PPs, then the expression "whereafter" constitutes, in miniature, a different grammar.[47] We leave this speculation as a suggestion which should be addressed in terms of a richer theory of speech register variation.

First-language acquisition

Now let us consider first-language acquisition from the perspective we have outlined. Stages in acquisition have always been seen as the movement from one grammar to another. However we have now argued that every speaker retains incompatible grammars. Therefore it is possible that a child retains an earlier stage when they move to a later stage. Why would a child retain multiple stages?

One answer could be that two social registers are involved. In other words, the earlier grammar has both a formal and a social definition. One can imagine that a child who has both "I want" and "me want" can express both a formal and a less formal kind of desire.

It is also a commonplace that children will treat a rule as optional which is later regarded as obligatory. For instance many children pass through a period in which inversion is optional:

[45] See Meisel and Ezeizabarrena (1996) for evidence that Baker's claims may not always hold.

[46] These are the formal options, but the reality is more complex. The presence of *passé composé* in some languages, but not others, may reflect the tense+verb option. However the reason why a language should move toward or away from this option is very obscure.

[47] There is more involved here than the syntax captures. We have: "therefore, thereof, therewith", where the unmoved form is completely disallowed in modern English: *"with there". The anaphoric property of "there" is maintained, but without the locative requirement (see Schafer and Roeper, 1999).

(64) (a) what he can do
(b) what can he do

The perspective advocated here would avoid the problem of stating optionality within a single grammar, which may be extremely difficult to do. If the WH-criterion (Rizzi, 1991) would mandate inversion, then why should it be optional in a child's grammar? Instead, we argue that the child actually retains two different grammars. De Villiers (1991) shows that children shift from non-inversion to inversion over several years, shifting each WH-word independently, as the child learns indirect question complementation for various verbs (''ask what he can do''). That is, ''what he can do'' shifts to ''what can he do'' two years before ''why he can sing'' shifts to ''why can he sing''.

In fact, (64a) might have a radically different structure, involving adjunction to IP or the generation, under Merger (Roeper, 1996), of a WH-word in the COMP position rather than the Spec of COMP. This generation of ''why'' under COMP continues to be present in the adult language:

(65) (a) why go downtown
(b) *where go downtown[48]

Thus the TB view leads naturally to the explanation of fairly subtle data in acquisition.

In addition, it provides an avenue to the most substantial puzzle in acquisition: why are stages less sharp than one would expect? Sudden shifts in grammar show that children use rules and not ''habits''. Thus Adam in the Brown Corpus suddenly uses 32 tags in one afternoon. However, there has always been evidence that children do not abandon previous structures at the moment they appear to adopt a new grammar. The Theoretical Bilingualism perspective may prove to be a very useful concept in this respect.

In sum, the customary view of acquisition is that the addition of a new feature to a grammar, such as a lexical item or a more abstract formal feature, simply deletes the previous representation. This remains a real possibility. A second avenue for development, however, is that the addition of a new feature changes the status of previous structures without entirely deleting them.

[48] Evidence that it is in the COMP position rather than Spec of Comp comes from the fact that long-distance movement is excluded:
i. why$_i$ say t_i [he can swim *t_i]
That is, the question is answered with why-say and not why-swim.

Summary

We have provided rather minute examples of where pockets of bilingualism may exist inside Standard English. We have discussed or mentioned isolated phenomena drawn from a variety of modules:

(66) (a) Case assignment
(b) resumptive elements
(c) do-insertion
(d) verb-final idioms
(e) WH- pre-posing in PP

In each instance we have argued that the generalization either follows principles of economy or remains lexically encapsulated.

Our sketch has arrived at a view of how Universal Grammar is deployed which constitutes a challenge to the common view of the consistency and uniformity of synchronic grammars, but is consistent with Chomsky's distinction between grammar and language. I have argued that Universal Grammar is available not only for the projection of wholly new L2 forms, but it is available within a given language to create radically different islands of grammar variation which in turn allow a nuanced array of communicative powers to the speaker.

We expect that as theory becomes sharper the pervasive presence of Theoretical Bilingualism within grammar will become more evident.

Real bilingualism

What has been under discussion is a kind of ''artificial bilingualism'' as seen from a quite technical perspective. It is quite obvious that real bilingualism is more intricate and complex. In addition there is a powerful phonological anchor which serves to separate two real languages. The speaker can assume that all rules linked to the phonology of one language do not, normally, penetrate another. Perhaps the microscopic interactions, at the lexical and social level, of ''artificial bilingualism'' will shed light on how different languages assume different social status (like registers) and how formal dissimilarities between two languages are represented within a single speaker.

References

Abdul-Kareem, L. (1996). The acquisition of case and ellipsis. Manuscript, University of Massachusetts.

Baker, M. C. (1988). *Incorporation: a theory of grammatical function-changing*. Chicago, London: University of Chicago Press.

Berwick, R. (1985). *The acquisition of syntactic knowledge*. Cambridge, MA, London: MIT Press.

Bickerton, D. (1975). *Dynamics of a creole system*. Cambridge: Cambridge University Press.

Bickerton, D. (1981). *Roots of language*. Ann Arbor: Karoma Press.

Budwig, N. (1989). The linguistic marking of agentivity and control. *Journal of Child Language*, 16, 263–284.

Cavar, D., & Wilder, C. (1996). Auxiliaries in Serbo-Croatian and English. Manuscript, University of Potsdam.

Collins, C. (1997). *Local economy*. Cambridge, MA: MIT Press.

Chomsky, N. (1986). *Knowledge of language*. New York: Praeger.

Chomsky, N. (1989, 1991). Some notes on economy of derivation and representation. In R. Freidin (ed.), *Principles and parameters in comparative grammar*, pp. 417–454. Cambridge, MA: MIT Press.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper and Row.

Davis, H. (1987). The acquisition of the English auxiliary system and its relation to lingiustic theory. Dissertation, University of British Columbia.

De Villiers, J. (1991). Why questions. In T. Maxfield & B. Plunkett (eds.), *Papers in the acquisition of WH*, pp. 155–173. University of Massachusetts Occasional Papers in Linguistics, Special Edition.

De Villiers, J., & Roeper, T. (1995). Barriers, binding and the acquisition of the NP-DP distinction. *Language Acquisition*, 4, 73–105.

Guéron, J., & Hoekstra, T. (1988). Les chaines T et les verbes auxiliares. *Lexique*, 7, 61–85.

Haegeman, L. (1990). Non-overt subjects in diary contexts. In J. Mascaro & M. Nespor (eds.), *Grammar in progress: GLOW essays for Henk van Riensdijk*, pp. 167–174. Dordrecht: Foris.

Håkansson, G. (1998). Language impairment and the realization of finiteness. In A. Greenhill, M. Hughes, H. Littlefield & H. Walsh (eds.), *Proceedings of the 22nd Annual Boston University Conference on Language Development*, pp. 314–324. Somerville, MA: Cascadilla Press.

Herschensohn, J. (1998). Minimally raising the verb issue. In A. Greenhill, M. Hughes, H. Littlefield & H. Walsh (eds.), *Proceedings of the 22nd Annual Boston University Conference on Language Development*, pp. 325–336. Somerville, MA: Cascadilla Press.

Hollebrandse, B., & Roeper, T. (1996). The concept of do-insertion and the theory of INFL in acquisition. In C. Koster & F. Wijnen (eds.), *Proceedings of the Groningen Assembly on Language Acquisition*, pp. 261–272. Den Haag: CIP-Gegevens Koninklijke Bibliotheek.

Holmberg, A. (1986). Word order and syntactic features in Scandinavian languages and English. Manuscript, University of Stockholm.

Van Kampen, J. (1996). PF/LF conversion in acquisition. In K. Kusumoto (ed.), *Proceedings of NELS 26*, University of Massachusetts: GLSA.

Kayne, R. (1994). *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Kroch, A., & Taylor, A. (1997). Verb movement in Old and Middle English: Dialect variation and language contact. In A. van Kemenade & N. Vincent (eds.), *Parameters of morphosyntactic change*, pp. 297–325. Cambridge: Cambridge University Press.

Labelle, M. (1990). Predication, WH-movement and the development of relative clauses. *Language Acquisition*, 1, 95–118.

Lebeaux, D. (1990). The grammatical nature of the acquisition process: Adjoin-a and the formation of relative clauses. In L. Frazier & J. de Villiers (eds.), *Language processing and language acquisition*, pp. 13–82. Dordrecht: Kluwer Academic Press.

Maxfield, T., & Plunkett, B. (eds.) (1991). *Papers in the acquisition of WH*. University of Massachusetts Occasional Papers in Linguistics, Special Edition.

Meisel, J. (1994). Getting FAT: Finiteness, Agreement and Tense in early grammars. In J. M. Meisel (ed.), *Bilingual first language acquisition: French and German grammatical development*, pp. 89–129. Amsterdam: Benjamins.

Meisel, J., & Ezeizabarrena, M. J. (1996). Subject-verb and object-verb agreement in early Basque. In H. Clahsen (ed.), *Generative perspectives on language acquisition: Empirical findings, theoretical considerations and cross-linguistic comparisons*, pp. 201–239. Amsterdam: Benjamins.

Müller, N. (1993). *Komplexe Sätze: Der Erwerb von COMP und von Worstellungsmustern bei bilingualen Kindern*. Tübingen: Gunter Narr Verlag.

Müller, N. (1998). Transfer in bilingual first language acquisition. *Bilingualism: Language and Cognition*, 1, 151–171.

Penner, Z. (1994). *Learning-theoretic perspectives on language disorders in childhood*. Manuscript, University of Bern.

Penner, Z., & Roeper, T. (1998). Trigger theory and the acquisition of complement idioms. In N. Dittmar & Z. Penner (eds.), *Issues in the theory of language acquisition: Essay in honor of Juergen Weissenborn*, pp. 77–111. Bern: Peter Lang.

Penner, Z., & Wymann, K. (1998). Normal and impaired language acquisition. Studies in lexical, syntactic and phonological development. Arbeitspapier 89.

Perez-Leroux, A. (1995). Resumptives in the acquisition of relative clauses. *Language Acquisition*, 4, 105–138.

Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.

Pollock, J. (1989). Verb movement, universal grammar and the structure of IP. *Linguistic Inquiry*, 20, 365–424.

Powers, S. (1996). Early subjects and agreement. Manuscript, University of Potsdam.

Randall, J. (1980). -ity: A study in word formation restrictions. *Journal of Psycholinguistic Research*, 9, 523–534.

Ravem, R. (1978). Two Norwegian children's acquisition of English syntax. In E. Hatch (ed.), *Second language*

acquisition: A book of readings, pp. 148–154. Rowley, MA: Newbury House.

Rizzi, L. (1991). Residual verb second and the WH-Criterion. In R. Clark, L. Rizzi & E. Wehrli (eds.), Technical reports in formal and computational linguistics no.2. University of Geneva.

Roeper, T. (1993). The least effort principle in child grammar: Choosing a marked parameter. In W. Abraham & E. Reuland (eds.), *Knowledge and language Vol. 1: From Orwell's problem to Plato's problem*, pp. 71–104. Dordrecht: Kluwer Academic Press.

Roeper, T. (1996). The role of merger theory and formal features in acquisition. In H. Clahsen (ed.), *Generative perspectives on language acquisitio: Empirical findings, theoretical considerations and crosslinguistic comparison*, pp. 415–449. Amsterdam: Benjamins.

Roeper, T., & de Villiers, J. (1992). The One Feature Hypothesis. Manuscript, University of Massachusetts.

Roeper, T., & Perez, A. (1997). The interpretation of bare nouns in semantics and syntax: Inherent possessors, pied-piping, and root infinitives. In J. Schaeffer (ed.), *Bare nouns and root infinitives*, pp. 153–175. MIT Working Papers in Linguistics.

Roeper,T., & Rohrbacher, B. (1994). Null subjects in early child English and the theory of economy of projection. Manuscript, University of Massachusetts.

Rubin, E. (1996). *Talk on bilingualism and minimalism*. Amherst, MA: University of Massachusetts.

Schafer, R., & Roeper, T. (1999). The role of the expletive in the acquisition of a discourse anaphor. Manuscript, University of Massachusetts.

Schütze, K., & Wexler, K. (1996). Subject case licensing and English root infinitives. In A. Stringfellow, D. Cahana-Amitay, E. Hughes & A. Zukowski (eds.), *Proceedings of the 20th Annual Boston University Conference on Language Development*, pp. 670–682. Somerville, MA: Cascadilla Press.

Speas, M. (1994). Null arguments in a theory of economy of projection. In E. Benedicto & J. Runner (eds.), *Functional projections*, pp. 179–208. University of Massachusetts Occasional Papers 17.

Stromswold, K. (1995). The acquisition of subject and object questions. *Language Acquisition*, 4, 5–48.

Takahashi, M. (1989). Object inversion in WH-questions. Manuscript, University of Massachusetts.

Tamanji, P. (1998). Verb movement: An alternative account. Manuscript, University of Massachusetts.

Vainikka, A. (1990). The status of grammatical default systems: Comments on Lebeaux. In L. Frazier & J. de Villiers (eds.), *Language processing and language acquisition*, pp. 83–103. Dordrecht: Kluwer Academic Press.

Vainikka, A. (1994). Case in the development of English syntax. *Language Acquisition*, 3, 257–325.

Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua*, 106, 23–79.

Wexler, K., & Culicover, P. (1980). *Formal principles of language acquisition*. Cambridge, MA: MIT Press.

Yang, C. (1999). The variational dynamics of natural language: Acquisition and change. Manuscript, MIT.

Zwart, J. (1993). *Dutch syntax: A minimalist approach*. PhD. dissertation, University of Groningen.

Received November 6, 1998 Revision accepted June 28, 1999

PEER COMMENTARIES

Bilingualism and the "mixed language" phenomenon

DALILA AYOUN
Department of French and Italian, University of Arizona, P.O. Box 210067, Tucson, AZ 85721, United States.
E-mail: *ayoun@u.arizona.edu*

Roeper (this volume) proposes that children are not only bilingual but also multilingual in the sense that they may entertain two or more grammars at once at different stages of development. This definition of bilingualism, or Theoretical Bilingualism, is significantly enlarged to include (i) adult speakers of a single language who routinely use different dialects or social registers of that language; (ii) child speakers of a single language who entertain several mini-grammars for different domains as they go through different stages of development.

I would like to extend Roeper's Theoretical Bilingualism to all adult speakers of so-called "mixed languages", i.e., languages which apparently exhibit both settings of parameters of Universal Grammar traditionally considered to be mutually exclusive, as opposed to (partially) inclusive or co-occurring. Although this suggestion may sound quite radical, it has already been mentioned in the literature, first by Atkinson (1990, p. 13), who noted that "nothing in principle rules out the possibility of multiple switch-settings", and second, by Fodor (1998, p. 21), according to whom:

> The two values of a parameter are standardly assumed to be mutually exclusive. This is not a necessary truth. Of course, no one construction can have both values ... but it does not follow that a language cannot have both values as options (it would be a superset language, subject to the Subset Principle). For example, Chomsky (1993) suggests that Arabic may have both strong and weak Tense features.

I will argue that verb movement phenomena in Arabic, binding phenomena in Dutch and Spanish, and null subject phenomena in Hebrew, all support the claim of mutually or (partially) inclusive parameter settings for different structures within the same language. If this argument is correct, then monolingual speakers of these "mixed languages" are actually "bilingual" speakers since their grammars entertain both parameter settings, which are otherwise instantiated separately in different languages.

Arabic and word order agreements

Standard Arabic uses two different word orders, SVO and VSO, along with different inflectional patterns. With an SVO order, the verb agrees in person, gender and number with a full NP subject. But with a VSO order, the verb agrees with its subject only in person and gender. Bolotin (1995, p. 20) suggests that the Arabic SVO/VSO alternation and asymmetrical agreement patterns are explained by a set of four inflectional parameters reproduced in (1):

(1) Inflectional parameters

	VSO order	SVO order
V features of T	strong	strong
V features of Agr	weak	strong
N features of T	weak	weak
N features of Agr	weak	strong

First, and following a minimalist account, the strong verbal features for Tense trigger verb movement to AGR for the VSO order, while the strong verbal and nominal features on Agr trigger verb movement to AGR and subject movement to [Spec, AGRP].

Second, "the agreement alternation that occurs between poor VSO agreement languages like Arabic and rich VSO agreement languages like Berber can be explained by assuming the following parameter values. All settings remain the same, except for the verbal features of AGR" (Bolotin, 1995, pp. 22–23) as reproduced in (2):

(2)

	poor VSO (Arabic)	rich VSO (Berber)
V features of T	strong	strong
V features of Agr	weak	strong
N features of T	weak	weak
N features of Agr	weak	weak

Thus the different parameter settings yield the following: the strong verbal features of AGR trigger verb movement to AGR as well as rich verbal agreement, while the weak verbal features of AGR exclude verb movement and verbal agreement is impoverished. To sum up, Bolotin (1995, p. 24) generalizes agreement patterns as in (3) and word orders as in (4):

(3) strong V features on AGR Arabic SVO; Berber VSO
weak V features on AGR Arabic VSO

(4) strong N features on AGR Arabic SVO
weak N features on AGR Arabic VSO; Berber VSO

Thus, strong features lead to overt verb movement, which then influences word order and agreement properties. Following Chomsky (1992), Bolotin (1995, p. 25) extends these parameter settings to adjectival agreement in Arabic, which presents similar asymmetries:

> For subject–adjective order, since both the verbal/adjectival features of AGR and the nominal features of AGR are strong, the adjective will raise to AGRAP and the subject to [Spec, AGRAP]. Once again, since the head (in this case, an adjective) is in AGR, rich agreement occurs. No such movement will occur for adjective subject order.

In other words, the strong and weak verbal features traditionally assumed to be mutually exclusive appear to be mutually inclusive or co-occurring in Arabic and Berber.

The idiosyncratic agreement pattern noted in Arabic occurs also in other languages: (i) in Breton, negative clauses show number agreement only in the VSO order; (ii) several dialects of Italian exhibit full agreement with SVO but only person agreement with VSO (Brandi and Cordin, 1987); (iii) in several dialects of Dutch and in Standard Dutch (second person singular only) the verb agrees with its subject in SVO but not VSO (Zwart, 1993). Bolotin (1995) concludes that a wide variety of languages, from Arabic and Berber to Welsh, Italian and Dutch, are best accounted for by a common parameter setting: weak features on AGR lead to poor agreement while strong features on AGR result in strong agreement. If this account is correct, these languages qualify as "mixed" languages and their speakers are bilingual in the sense that they use both strong and weak Tense.

Binding theory

Koster (1994) applies this concept of strong versus weak morphological features and Checking Theory to anaphoric chains to reformulate Principle B of the Binding Theory. Principle B predicts that bound pronouns should not be found in local domains, but this prediction does not seem to hold in several languages including English and Dutch. The culprit appears to be the morphological distinction of anaphors and pronouns. Koster proposes to redefine anaphors as locally bound NPs, so that both *him* and *himself* may be anaphors. In Koster's analysis, there are "two kinds of morphological distinction: (i) specialized forms (like Dutch "zich") versus nonspecialized forms (like English "him"); and (ii) short forms ("zich", "him") versus long forms ("zichself", "himself")" (p. 45). Anaphors have strong morphological features which must be checked in the appropriate SPEC-head configuration, i.e., a strong head. This requirement introduces a parametric distinction between languages: to be licensed, morphological anaphors must agree with one of the two functional nodes, AGR-S or AGR-O. This parametric distinction is as follows: (i) languages such as Frisian do not need anaphoric agreement; (ii) languages such as German and Slavic select AGR-S; (iii) languages such as French and English select AGR-O. Moreover,

> Reality is somewhat more complex in the sense that certain languages, like Dutch and Spanish, seem to have both possibilities. Such languages select either the German-Slavic option (AGR-S) or the English-French option (AGR-O). ... Dutch can choose between two distinct grammars in this respect. (Koster, 1994, p. 49)

In other words, "Dutch may select either of the systems allowed by the possible parameter settings" (p. 56), and thus represents another example of mutually inclusive, as opposed to co-occurring, parameter settings. Koster's examples reproduced here in (5) illustrate the standard Dutch AGR-S option, while the examples in (6) show that different anaphoric forms can exploit the AGR-O option:

(5) (a) *Jan* wast *zichzelf*
"John washes himself"
(b) *Jan* sprak over *zichzelf*
"John talked about himself"
(c) *Jan* zag een slang naast *zich*
"John saw a snake next to him"

(6) (a) **Jan* wast *hemzelf*
(b) *Jan* sprak over *hemzelf*
(c) *Jan* zag een slang naast *hem*

Koster points out that although the examples in (5) are clearly more standard than those in (6), "Dutch uses *either one of two* parameter settings as exemplified by Spanish as well" (p. 44, emphasis mine):

(7) (a) *Juan se* lava
John self washes
"John washes"
(b) *Juan* habla de *s' mismo/_l (mismo)*
"John talks about himself"
(c) *Juan* vió una serpiente junto a *sí* mismo/ él (mismo)
"John saw a snake near him (self)"

If Koster's proposed AGR-S/AGR-O parameter receives additional cross-linguistic support, it will be another example of a parameter which may be instantiated with both settings in mixed languages.

Null subject phenomena

Berman (1990, p. 1141) argues that Hebrew is a "mixed" language as well when it comes to the syntactic phenomena subsumed under the null subject parameter:

> Like Russian, [Hebrew] disallows a 3rd person impersonal pronoun comparable with English "they" or "one", French "on" or German "man" in plural impersonals like [8]; but it may have a pleonastic "it" subject in non-referential impersonal constructions like [9] and [10]. If a language tolerates both expletive and null subjects, expletives will occur in the pro-sentential function noted here for Hebrew rather than in existential contexts like English "there". A strictly subject-requiring language like English will demand a pleonastic subject in both environments; a language like Hebrew that is mixed with respect to zero subjects will allow pleonastics that are syntactically anaphoric (like Hebrew "ze"); and a uniformly null-subject language like Italian will disallow any kind of pleonastic subject.

Berman's examples are replicated here:

(8) (a) Lo ovd-im be-shabat ba-arets
not work-PL on-Saturday in-Israel
"they/you/people don't/one doesn't work on a Saturday in Israel"
(b) Ya'avd-u sham be-mishmarot
will-work-PL there in-shifts
"They'll work on shifts there"

(9) (a) Haya nora kar sham
was very cold there
(b) Haya li mesha'amem ito
was to-me boring with-him
"I was bored in his company"

Colloquial Hebrew also allows the pleonastic or expletive "ze" as illustrated in the following conversation between

a mother and her two-year-old as they look at a storybook:

(10) Child: ma hu ose la?
"What he does to-her?"
Mother: o marbits la makot. O moshex la base'arot. *Ze* yafe kaxa lariv?
"Gives her hits. Pulls to-her at-hair. It nice so to-fight?"
Child: *ze* lo zafe
"it not nice"
Mother: al ma, al ma hi tsiyra, ha-buba?
"On what on what she drew the doll"
ma se? nyar? lo, ze ha-kir
"what's that? Paper? No, it's the wall"
o mutar le-tasyer al ha-kir?
"(is-it) allowed to-draw on the-wall"
mi tsiyer al ha-kir shelanu. Sivani tsiyra. Ve *ze* yafe?
"who drew on our wall? Sivani drew. And (is) it nice?"
Child: o lo yafe bixlal!
"Not nice at all"

Berman further observes that null subjects are not consistently and uniformly omitted: (i) in the past tense, subjects are generally omitted only for the 1st and 2nd person, singular and plural; (ii) in the future tense, the 1st person singular pronoun "ani" is never omitted and the 2nd person singular pronouns, "at", "ata", and plural "atem" are rare. It thus appears that the licensing of null subjects in the past and future tenses "interacts with variables of morphophonological, semantic, and pragmatic distinctiveness, as well as distinguishing neutral from contrastive statements" (Berman, 1990, pp. 1145–6). In spite of this complexity, children acquire Hebrew without difficulty and end up with a "bilingual" grammar with both overt and null subjects.

Conclusion

I have proposed that Roeper's Theoretical Bilingualism be extended to speakers of "mixed languages" defined as languages allowing (partially) inclusive or co-occurring parameter settings for different structures, as shown in Koster (1994) and Berman (1990) for example. This proposal implies a permanent grammatical bilingualism among adults in addition to the permanent social bilingualism among adults or the temporary grammatical bi- or multilingualism among children suggested by Roeper. Bilingualism, and multilingualism, would then be the norm, not the exception (Cook, 1993). The suggestion of bilingual speakers for mixed languages is also reminiscent of the concept of the simultaneous operation of two grammars as opposed to having one language active to the exclusion of the other in "true" bilinguals as argued by Muysken (1998, p. 200):

> in some cases of code-mixing, components of both languages are not active in sequence, but simultaneously. The main evidence from simultaneous representations comes from mixed verbal complexes and from the phenomenon of delayed lexicalization, the non-linear interaction of the two grammars.

References

Abd El-Moneim, A. (1989). The role of INFL. Doctoral dissertation, University of Connecticut, Storrs.

Atkinson, M. (1990). The logical problem of language acquisition: Representational and procedural issues. In I. Roca (ed.), *Logical issues in language acquisition*, pp. 1–31. Dordrecht: Foris.

Berman, R. (1990). On acquiring an (S)VO language: Subjectless sentences in children's Hebrew. *Linguistics*, 28, 1135–1166.

Bolotin, N. (1995). Arabic and parametric VSO agreement. In M. Eid (ed.), *Perspectives on Arabic linguistics VII,* pp. 7–27. Amsterdam: John Benjamins.

Brandi, L., & Cordin, P. (1987). Two Italian dialects and the null subject parameter. In O. Jaeggli & K. Safir (eds.), *The null subject parameter,* pp. 111–142. Dordrecht: Kluwer.

Chomsky, N. (1992). A minimalist program for linguistic theory. MIT Occasional Papers in Linguistics 1.

Chomsky, N. (1993). A minimalist program for linguistic theory. In K. Hale & S. J. Keyser (eds.), *The View from Building 20: Essays in Honor of Sylvain Bromberger*, pp. 1–52. Cambridge: MIT Press (reprinted in Chomsky, N. (1995)).

Cook, V. J. (1993). *Linguistics and second language acquisition*. London: Macmillan.

Fassi Fehri, A. (1988). Agreement in Arabic, binding and coreference. In M. Barlow & C. Ferguson (eds.), *Agreement in natural language,* pp. 107–158. Stanford: Center for the Study of Language and Information.

Fodor, J. D. (1998). Unambiguous triggers. *Linguistic Inquiry*, 29, 1–36.

Koster, J. (1994). Toward a new theory of anaphoric binding. In B. Lust, M. Suñer & G. Hermon (eds.), *Syntactic theory and first language acquisition: Crosslinguistic perspectives*, pp. 41–69. Hillsdale: Lawrence Erlbaum.

Mohammad, M. (1990). The Problem of subject-verb agreement in Arabic: Towards a solution. In M. Eid (ed.), *Papers from the First Annual Symposium on Arabic Linguistics. Perspectives on Arabic Linguistics*, Vol. I, pp. 95–125. Amsterdam: Benjamins.

Muysken, P. (1998). We are all native speakers, but of which language? In R. Singh (ed.), *The native speaker: Multilingual perspectives*, pp. 193–204. New Delhi: Sage Publications.

Zwart, J.-W. (1993). Verb movement and complementizer agreement. In J. Bobajik & C. Phillips (eds.), *Papers on case and agreement I*, pp. 297–340. Cambridge, MA: MIT Press.

 Bilingualism: Language and Cognition 2 (3), 1999, 190–192

A blueprint or a puzzle?

HUBERT HAIDER
University of Salzburg, Muehlbacherhofweg 6, 5020 Salzburg, Austria.
E-mail: *hubert.haider@sbg.ac.at*

Questioning the background assumptions

Current versions of Generative Grammar take for granted two implicit premises that are in fact worth being questioned, namely that (i) the language faculty is a cognitive sentence generation machinery (''numeration'' as a selection from the inventory of basic units; operations on the inventory, as ''merge'', ''move'', ''check''); and that (ii) the grammar acquisition process of the child is *guided by UG as a universal blueprint* for sentence structuring. Premise (i) rules out optionality on the level of rules, and premise (ii) suggests non-optionality on the level of acquisition, that is, a cross-individually uniform progression. *Universal bilingualism* defends these premises against conflicting data. What I will suggest is a more radical move: the conflicts are apparent because the premises are inadequate.

The first premise is potentially ill-founded for the following reason: It is the receptive aspect (*string-to-structure mapping*), and not the perspective of generation (*combinatorics of elementary units*), that is the decisive factor, both in on-line usage, in the course of grammar acquisition and in an evolutionary perspective. It is a truism that the specifics of a grammar enter the head of the learner only if they are processable. The learner is a cognitive filter that selects. So, first of all, the cognitive constitution of the learner determines what is a possible grammar. Secondly, the learner is a recipient. The on-line task of the learner and in fact any listener is to map a one-dimensional array of terminals (= PF) on to an at least two dimensional representation, namely *semantic form* (= SF). I use the theory-neutral term SF, rather than LF. A representation of SF is what is modeled for instance by a DRS in the Discourse Representation Theory, that is, a hierarchically structured, complex representation accessible to a model-theoretically founded deductive system. The grammar of a language L is the recursive definition of the algorithm for this mapping in L. PF-to-SF mapping should be modeled with an adequate algorithm, that is, an algorithm that takes *linear arrays* as the input and assigns structures as output values. These structures must satisfy the interface requirements at both interfaces (PF- and SF-interface).

The *solution of the projection problem*, that is, the characterization of the general theory that selects the adequate grammar for L on the basis of a sufficiently diverse though limited subset of expressions of L is the fundamental theoretical problem. Hence UG should be viewed as the constraint space that guarantees the effective identification of the mapping algorithm from the perspective of the learner and listener.

The view sketched above is not merely a matter of perspective. It has distinct empirical implications. For instance, some problems of one approach cease to be problems for the other approach, for example the *optionality problem*. Optionality is a genuine problem in the sentence generation machinery-perspective, for a simple reason: why should the machinery continue generating a derivation with an extra step if the product is well formed already? If grammar is viewed as a projection algorithm, however, there is no optionality problem of this kind. A given array of terminals is mapped on the minimal convergent projection:

(1) (a) Would only those *who appreciate this problem* agree?
    (b) Would only those agree *who appreciate this problem*?

The difference between (1a) and (1b) is usually described as the result of *optional* extraposition of the relative clause. In the projection approach, (1a) and (1b) are two different arrays of terminals that are mapped on different, convergent projections. True ''optionality'' is a systematic relation between SF and PF, namely the converse of the ambiguity relation. If the mapping of a given PF-array to SF is not unique, this array is called (structurally) ambiguous. On the other hand, if the mapping of a given SF to PF is *systematically ambiguous*, this is (structural) *optionality*. Since grammars do not embody absolute constraints against ambiguity in one direction (i.e. PF-to-SF), there is no theory-independent reason for worries about ambiguity in the converse direction (i.e. SF-to-PF). Optionality is expected, for instance, if the construction space is underdetermined (see example 2; details in Haider, 1997):

(2) (a) She insisted that he *not have* left before I return
    (b) She insisted that he *have not* left before I return

Finite auxiliaries raise to a functional head position above negation in English. In the so-called subjunctive construction (2), the finite verb does not show subject agreement. As a finite verb, it may substitute for a functional head, but there is no agreement trigger. Underdetermination of the construction space may result in partial indeterminacy in convergent structures.

Underdetermination is also the key concept for a proper understanding of the interim grammars of L1-acquisition: if a child's set of utterances is inconsistent with respect to the adult grammar that does not mean that the child's interim grammar is inconsistent; its construction space is less constrained than that of the adult grammar. Optionality may arise due to the underspecification of constraints of the adult grammar. The more data the child integrates, the tighter the construction space gets, converging ultimately to the input grammar.

The second premise (*UG as a blueprint for grammar assembling*) seeks to eliminate *optionality* on the level of explanatory adequacy: L1 acquisition is seen as a universally determined succession of interim grammars that converge to the target grammar. Cross-individual optionality in terms of distinct alternative succession paths are unexpected. In this picture, there is neither room for the case of *optionality* mentioned above, namely interim grammars with co-existing structures that do not co-occur in the target grammar, nor for optionality in terms of alternative acquisition paths. There are good reasons, however, to give up the second premise and replace the *blueprint-strategy* by a *puzzle-strategy:* You start a puzzle with whatever piece you choose, but the more pieces are integrated into larger units, the smaller the degree of freedom for integrating the remaining pieces becomes. Trivially, for the final piece there is just one position left. Children's behavior – on the evidence of large and diverse corpora of acquisition data (see e.g. Gawlitzek-Maiwald, Tracy and Fritzenschaft, 1992) is more likely to be "puzzling" than "blueprinting". In the earlier stages of grammar acquisition, the construction space is much less constrained (within the limits of UG) that at an advanced stage. The more "pieces of grammar" are integrated the more constrained is the construction space. The puzzle-strategy is globally indeterministic, but locally deterministic.

The UG-enforced overall consistency demands guarantee the convergence of the acquisition sequence to the target grammar. As in the case of a puzzle, neither the starting point is *intrinsically* determined, nor the continuation. There is room for much individual variation. And, what is crucial, structures may coexist at a given interim stage that do not coexist at a later stage. They co-exist if in comparison with the target grammar the construction space is under-specified and therefore does not exclude one of them or both. *Underdetermination* is the source of this kind of *optionality*.

n-lingualism or puzzle & periphery?

The hypothesis of "universal bilingualism" (UB) is intended primarily as an answer to questions that relate to optionality and variation in first language acquisition: "The notion of optionality [of rules] can be eliminated". But, in fact, the UB-hypothesis implicitly presupposes constant code-switching and thereby introduces optionality again. Instead of an *optional rule*, an alternation arises as the result of *optionally* switching *grammars*. So, the learnability problems remain in the form of eliminating "code-switching" and "particular grammars" in the course of language acquisition and the task remains intractable.

The existence of "particular grammars" is a widely acknowledged fact, in a different terminology, though. It is covered by the distinction between core and periphery (Chomsky 1981, pp. 8ff.). A "periphery"-phenomenon by definition is not consistent with the core grammar of a given language. Therefore a periphery element has its own local partial grammar. Take for instance the exceptional serialization pattern of "enough" and its Germanic cognates. In all Germanic languages they follow the predicate they modify.

(3) (a) This is clear *enough* / "*sufficiently* clear"
(b) Das ist klar *genug* / "*genügend* klar"

The role of potentially diverging partitionings of a language in core and periphery data for language change is analyzed by Andersen (1973). He introduces the notion of "via-rules" as links to non-core subsystems that guarantee that the different grammars of the speech community remain weakly equivalent in their unstable equilibrium. Periphery properties can be associated with lexical items as in (3), just as well as with construction types. V2-structures in English quotation environments are instances of a peripheral V2-construction in a non-V2 language. The case of "It is I" vs. "It is me" is another case of core – periphery shifts.

Periphery – this should not be overlooked – is a residue modulo consistency. Periphery phenomena are defined ex negativo: "not covered by core grammar". So in language acquisition, depending on the vantage point, a periphery either does not exist at all, or most of language acquisition is equivalent to integrating periphery (with respect to a given interim grammar of the child) into core grammar by extending the interim grammar. In other words, either there is no periphery for the child grammar, or everything new is periphery for the child grammar and needs to be turned into a core grammar instance. So far, acknowledging a core vs. periphery distinction seems to be sufficient. A child cannot know in advance what kind of phenomenon belongs to the core and the periphery, respectively, if one presupposes the puzzle strategy. Therefore, construction variants may coexist for a while. In the course of acquisition, degrees of freedom are eliminated and thereby some variants become obsolete in the child's grammar and disappear.

Dealing with details

The following phenomena are discussed as sources of evidence for UB: case assignment to subjects, the use of resumptive pronouns in relative clauses, do-insertion, verb-final idioms, and wh-preposing in PPs. I sketch briefly what I perceive as the null-hypothesis.

Case distinctions are hard to identify for the learner in English. As morphologically manifest properties of DPs, they can be observed only with a few pronouns. So if a child assumes that the distinction between "me" and "I" for subjects is a difference between a weak and a strong pronoun, the child entertains a viable UG-option.

Resumptive pronouns in relative clauses are cross-linguistically well established. The *such-that* construction frequently used in Montagovian contexts is a construction of formal English that illustrates this possibility. For the child, the use of a resumptive pronoun obviates the need of postulating an empty wh-operator that binds a trace and precedes the complementizer "that". It is a less complex construction in terms of the projection of structure, favored in earlier phases.

Do-insertion is the reflex of a single, exceptional prop-

erty of English: main verbs do not move to functional head positions. The discussion on feature weakness based on the evidence of the interpretation of the English present lacks force. The English grammatical present does not have the interpretation of the English progressive present (unlike German), simply because English (but not German) has a grammatical progressive present. Gricean implicatures suffice to derive this result. "Do" is an expletive auxiliary that appears in contexts that lack a local relation of a functional head feature and the VP. If the VP is immediately dominated by the projection of the functional head whose features need to be spelled out on the verb, the features may be lowered. This is impossible, if the very F-head has been raised (i.e. in I-to-C contexts) or if there is an intervening head (i.e. the case of the negation head "not"). In the latter case, the lowered features would end up on the negation element as the head of the complement of the functional head. It is not a *tense* but rather an *agreement* feature that is the triggering feature, as evidenced by the lack of *do*-support in negated "subjunctive" clauses, that is, tensed clauses without agreement. It is difficult to feel an urgent need for the assumption of hidden bilingualism in this respect. English do-support is the grammaticalization of a former light-verb construction: there are colloquial varieties of German and other Germanic languages, in which finite $do_{fin}+V_{inf}$ is generally in free variation with V_{fin} for activity verbs.

As for the specific case of wh-preposing with PPs, "where+P" is a good illustration of a periphery feature that is a lexicalized historic relic. "Where+P" – as in other Germanic languages – is the wh-counterpart of "there+P" (e.g. "thereafter, therefore, therein, thereto, thereunder", etc.), with "there" as a cliticized weak proform in the predecessor construction.

In sum

I consider the question still open as to whether the mild defects of grammar called periphery need to be cured with such a strong antidote as universal bilingualism.

References

Andersen, H. (1973). Abductive and deductive change. *Language*, 49, 567–595.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Gawlitzek-Maiwald, I., Tracy, R. & Fritzenschaft, A. (1992). Language acquisition and competing linguistic representations: The child as arbiter. In J. Meisel (ed.), *The acquisition of verb placement: Functional categories and V2 phenomena in language acquisition*, pp. 139–179. Dordrecht: Kluwer.

Haider, H. (1997). Economy in syntax is projective economy. In C. Wilder, H.-M. Gärtner & M. Bierwisch (eds.), *The role of economy principles in linguistics. Studia Grammatica 40*, pp. 205–226.

Are L2 speakers bilingual or bilexical?

ROGER HAWKINS
Department of Language and Linguistics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK.
E-mail: *roghawk@essex.ac.uk*

Few second-language researchers would deny that there is some intuitive content to the notions of "stages" and "optionality" in second language (L2) syntactic development. Some syntactic properties are used productively in early L2 grammars to a much greater degree than others, which appear later; many properties which are categorical in the target language are not initially used consistently, but alternate with zero or some other property. Trying to provide an explicit account of the intuitive content of these notions has been of considerable interest to L2 researchers. In the 70s and 80s the focus was on stages, in the 90s it has been on optionality.

Roeper's "theoretical bilingualism" proposal is an attempt to capture the intuition in a way which spans both L1 and L2 development, and even extends to mature native "monolingual" competence. By its intended breadth it expresses the strongest of "strong continuity" hypotheses. The idea, paradoxically, eliminates both "stage" and "optionality" from the domain of grammatical knowledge, treating them as epiphenomena. Optionality arises not because competence allows UG-sanctioned variants to coexist in the same grammar, but because competence consists of multiple grammars, each with its own categorical choice of UG options; and under different circumstances speakers may opt for one grammar or another. Stages are nothing more than changes in the balance of the lexical productivity of one grammar over another: at time *x* grammar *g* is used with a wider range of lexical classes than grammar *h*, but at time *y* the balance changes.

From the point of view of learnability, the proposal that grammatical operations are categorical is welcome. And some of Roeper's intriguing insights appear to open up new lines of possible enquiry into stages and optionality in L2. However, even in this avowedly programmatic initial sketch there appear also to be some problems. I will comment briefly on both, starting with the problems.

Given the minimalist perspective that Roeper assumes, a conceptual problem arises with regard to where UG-licensed syntactic variation resides. Roeper appears to assume that it is in the syntactic–computational component, hence the need to propose multiple grammars to handle variation within the same speaker. For example, the grammar of English does not generally allow verb-second (*"The witch shrieked not"), but in quotation contexts verb-second is an option ("'Go', shrieked the witch"). This, according to Roeper, is an effect of speakers of English having two grammars: one involving V2 which is applied to quotation contexts, the other involving non-V2 which is applied in most other declarative contexts. However, in recent work within the principles and parameters framework, it has been more and more assumed that the syntactic–computational component is invariant and that parametric variation is associated with the features of functional categories (Borer, 1984; Fukui, 1986; Chomsky, 1995, pp. 378–379). On this view V2 is the result of a particular feature specification of C which attracts thematic verbs in tensed clauses. If a language allows both non-V2 and V2, this must be the result of there being more than one C (or more than one specification of C) in the lexicon. V2 and non-V2 derivations would then be the effect of starting with arrays of lexical items which differ in the particular C selected. Crucially, the syntactic–computational component does not vary – and indeed the most "economical" assumption might be that it is universally invariant, the syntax of languages differing only where feature specifications of functional categories differ. If this is correct, the possibility of both V2 and non-V2 derivations in English is not the effect of English speakers having different grammars, but of having different lexical items belonging to the category C, their different distributions determined by different selectional properties (just as in recent work Chomsky, 1998, proposes that there are two T's in English: T with full tense and phi-feature properties, and "defective T", which occurs in Exceptional Case Marking contexts). Given this view, it might be more appropriate to describe speakers of English as "bilexical" rather than bilingual.

A second more specific problem concerns the account of the difference between thematic verb raising languages like German and French, and non-verb-raising languages like English. Roeper observes that thematic verb raising appears to have a semantic effect. For example, "Hans singt" has two interpretations: "Hans sings (habitually)" and "Hans is (at the moment) singing". English "John sings" only allows the first of these interpretations. Roeper argues that overt verb raising has the effect of "anchoring" the event described by the verb to the context of utterance, giving rise to the "present tense" interpretation. "Sings" does not raise in English, is not anchored to the context of utterance, and so only has a "generic" event interpretation. The observation seems to be correct. Roeper then links this to a proposal that in German and other thematic verb-raising languages, there is a strong affix belonging to the category T, to which the verb raises. This strong affix effectively links the verb to the tense property of T. In English, the 3rd person *-s* is a weak affix which attaches to the verb in the lexicon, and does not link directly with the tense property of T. However, with negation and in interrogatives a strong affix appears, and is "spelled out" via do-insertion.

However, the idea that affixes might be located under T, with verbs raising to T to form a morphologically complex

word, seems to be a step backwards in terms of minimalist assumptions. Chomsky has suggested (1995, p. 239) that "on the simplest assumptions, the lexical entry provides, once and for all, the information required for further computations – in particular, for the operations of the phonological component (including morphology ...)". Thus, whatever the mechanism that forces the verb to raise, it is implausible that it is a morphological affix. Moreover, treating *do* as the spell-out of a strong affix gives the wrong interpretive result. If the semantic function of a strong affix under T is to anchor an event to the context of utterance, the prediction would be that when "do" is present, the "present tense" interpretation is possible:

(1) (a) John doesn't sing
    (b) Does John sing?

But it seems that (1a) cannot be interpreted as "John isn't singing" and (1b) cannot be interpreted as "Is John singing?".

A more plausible option for explaining verb raising arises as a consequence of one of Roeper's insights. He suggests (echoing ideas in Chomsky (1998)) that "all movement is accompanied by some semantic distinction". This certainly seems to be correct in the case of thematic verb raising; the verb raises to create a specific interpretation ("tense anchoring"). It is unnecessary to propose that there is an affix located under T or even that there are the rather unexplanatory "strong features" present in T. If this idea can be maintained more generally for cases where there is movement of lexical categories to functional categories, the notion of uninterpretable features driving movement would be unnecessary. Functional categories which force movement would "mean something different" from functional categories which do not. This would be consistent with the idea that syntactic variation is determined in the lexicon.

Consider now development and stages. Roeper's insight is that development results from a tension between different sorts of economy principle. In the earliest stages of L1 acquisition children prefer derivations which obey UG-sanctioned economy (shortest move, checking under direct c-command, and so on) – the preferred options are referred to as "minimal default". But this preference soon comes into conflict with principles of local economy in the particular language they are exposed to, or other principles like "meaning explicitness". The tension between these causes minimal default preferences to coexist with more target-like preferences. For example, Roeper observes that L1 learners of English, German and Dutch opt for do-insertion in cases not permitted by the target language because do-insertion allows checking under direct c-command. But these cases often alternate with derivations which do not have do-insertion.

How are these ideas interesting for the investigation of stages and optionality in L2 syntax? Consider the following case. In one well-known set of studies, White (1990/91, 1991, 1992) has shown that French-speaking adolescent learners of English in Canada do not typically transfer thematic verb raising into English negative and interrogative contexts. They show a strong preference for do-insertion in both contexts (on a preference task they chose do-insertion in over 80 per cent of cases, and of the 1171 questions they produced only 0.17 per cent involved thematic verb raising). By contrast, with verb phrase adverbs they persisted in using French verb raising alongside English non-raising in equal measure, even after exposure to considerable positive evidence for the English case. Such behaviour might be explicable on the basis of Roeper's insights. Even if these L2 learners initially transfer the specification of French functional categories into their lexicons for English, once they encounter do-insertion, by economy principles they should prefer this both over English non-raising, and even over French raising, because merger of a free form with T allows feature checking under direct c-command, which is possible neither with verb raising (since we have abandoned the idea that there is a strong affix in T) nor non-raising. The optionality of verb raising in other contexts might persist if learners establish an English T alongside transferred French T in their lexicons. They would then be bilexical as far as T is concerned. Economy considerations do not arise here, since raising versus non-raising has different effects on meaning. The problem for French learners then is determining the lexical range of English T.

References

Borer, H. (1984). *Parametric syntax*. Dordrecht: Foris

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, N. (1998). Minimalist inquiries. *MIT Occasional Papers in Linguistics, 15.*

Fukui, N. (1986). A theory of category projection and its applications. Unpublished doctoral dissertation, MIT.

White, L. (1990/91). The verb movement parameter in second language acquisition. *Language Acquisition*, 1, 337–360.

White, L. (1991). Adverb placement in second language acquisition: Some effects of negative evidence in the classroom. *Second Language Research*, 7, 133–161.

White, L. (1992). Long and short verb movement in second language acquisition. *Canadian Journal of Linguistics*, 37, 273–286.

How universal is "theoretical bilingualism"?

AAFKE HULK
Romance linguistics/HIL, University of Amsterdam, Spuistraat 134, 1012 VB Amsterdam, Netherlands.
E-mail: *Aafke.hulk@hum.uva.nl*

In Universal Bilingualism Roeper makes the interesting, strong claim that his concept of Theoretical Bilingualism allows one to generalize over all cases where two properties exist in a language that are not statable within a single grammar. This holds not only for developmental optionality and stages in L1 (and L2) acquisition, but also for different, stable domains within the (apparently monolingual) adult grammar. In his view principles of Economy of Representation and of Meaning Expliciteness compete to explain the coexistence of two or more variants of a given construction.

In this short note I would like to raise the question whether Roeper's theory really allows us to account for "mature/stable" and "developmental" optionality by the same universal principles. This question also relates to the "trigger" problem: how come some grammars may coexist for a long time (as in different registers and in true bilinguals), whereas other grammars only coexist briefly and develop into new ones (as in acquisition and language change). I will briefly review what the universal principles proposed by Roeper tell us about these questions.

In the course of language development, children go through stages of temporary optionality which allow for the co-existence of forms that are incompatible in adult grammar. During the transition phases between stages the child will show features of both the old and the new grammar. Eventually children abandon this type of optionality and retain the target option from the adult grammar. In Roeper's view, all children go through three stages in which the two guiding principles of his theory compete. In the first stage Economy of Representation outranks Meaning Expliciteness and in subsequent stages Meaning Expliciteness outranks Economy of Representation. Apparently there is some universal driving force or trigger that makes the children re- rank these principles. Roeper does not develop this point. Moreover, it is unclear which role these principles play in the case of optionality found in mature grammars.

As a universal phenomenon, Roeper also mentions the role of context: "if one grammar permits subject-drop and the subject is contextually clear, then this circumstance may influence the choice of grammar".

According to Roeper, this option holds not only for children, but also for the adult who controls several dialects and for the true bilingual, who selects Spanish or English on different occasions. It seems rather unlikely, however, that code-switching between languages is mainly determined by such things as the contextual licensing of the (null) subject of the clause.

Elsewhere Roeper assumes that the principle of Economy can be overruled by specific lexical cases, for example in the residual V2 cases in (adult) English, creating apparent optionality of V-movement. The existence of optionality within adult grammar has been extensively discussed in the literature. To mention just a few examples: scrambling in West-Germanic, (wh-)questions in French, embedded verb clusters in Dutch. Optionality phenomena pose a challenge to formal grammatical theories, such as Minimalism and Optimality Theory. There are several ways to allow for optionality within an economy-based theory: the idea that it reflects the existence of different (parts of) grammars, as adopted by Roeper and others, may offer a solution. The consequence of this idea is that monolinguals are in fact multi-linguals, but since "real" multilingualism is the regular situation for many or most human beings, this is not so exceptional as it may seem. We also know that adult speakers generally can switch freely between all their grammars/languages. Roeper attributes these switches to social factors, involving among other things speech registers.

Again it is not yet entirely clear what the role of social factors is in "the other type of universal bilingualism", developmental optionality. Roeper proposes that different grammars can be localized either in lexical classes or in speech registers. This raises the question whether this is the only explanation for the different grammars in child language. The same question holds for interlanguage grammars in L2 acquisition which also show developmental optionality: just as in L1 acquisition L2 learners go through different stages towards the target-grammar. There is however an important difference too: whereas many linguists assume that children start their L1 acquisition with some form of default setting for the principles to be acquired, this is not the general view of the starting point of L2 acquisition. There the role of the L1 grammar, especially in the initial stages, is very important and moreover it is an extra source for optionality.

Finally, Roeper's idea of Universal Bilingualism offers interesting perspectives on a parallel treatment of monolingual and (true) bilingual language development. Let us briefly look at the acquisition of syntax by children who are balanced bilinguals, exposed to two languages from birth. Recent work has shown that, although these children separate the two languages from very early on, there is nonetheless evidence of cross-linguistic influence in syntax. The interesting question is to find out which parts of grammar are sensitive to such influence and why this should be so. Müller, Hulk and Jakubowicz (1998) argue that crosslinguistic influence is expected to be possible in exactly those areas which are also problematic, to a lesser extent, for monolingual children. They discuss data from bilingual French/Dutch, French/German and Italian/

German children which show influence of the Germanic language on to the Romance language in the case of object drop. In this case, the input in the Romance language allows for more than one analysis and the input in the Germanic language contains positive evidence for one of the possible analyses, creating confusion and delay in the acquisition process. Moreover, they argue that this indirect influence is possible because the phenomenon in question concerns the interface between two modules of grammar: discourse and syntax. The bilingual children are confronted with the interaction between a universal discourse strategy, language specific morpho-syntactic licensing rules and a partially overlapping input. What can Roeper's Theoretical Bilingualism tell us about such phenomena? His idea that Economy of Representation outranks Meaning Explicitness in the early stages of language development could probably explain that we find object drop (as topic drop) in these early stages of all children. It is not immediately clear how this relates to the differences in object drop between Germanic and Romance children. It is not yet clear either what this can tell us about the way in which the acquisition by bilinguals is different from the acquisition by monolinguals and where we expect crosslinguistic syntactic influence.

Summarizing, in two respects developmental optionality is crucially different from (stable) optionality found in adults: (i) it is a transitional phenomenon, and (ii) the coexistence of different constructions cannot be attributed (solely) to different speech registers or subtle semantic differences. Roeper's Theoretical Bilingualism offers us a number of elements with which we may eventually develop an overall picture of both types of bilingualism. However, some of the proposed principles seem to be particularly relevant for developmental optionality only, whereas others seem to play mainly in mature bilingualism. The ''real'' bilinguals – both children and adults – are yet another group to be taken into account. Roeper's Theoretical Bilingualism makes a strong and interesting claim, suggests bits of solutions, but above all, it raises a lot of yet unanswered questions concerning the possibility of a unique account of Universal Bilnguialism.

References

Müller, N., Hulk, A., & Jakubowicz, C. (1999). Object omissions in bilingual children: Evidence for crosslinguistic influence. In A. Greenhill, H. Littlefield & C. Tana (eds.),*Proceedings of the 23rd Annual Boston University Conference on Language Development*, pp. 482–494. Somerville, MA: Cascadilla Press.

Bilingualism or inherent variation?

MARJORY MEECHAN
c/o ZTI, Box 369, Horizon House, McGregor St., Bridgetown, Barbados.
E-mail: *meechan@yahoo.com*

Given that variability is found in every language in the world, whether spoken, written or signed, either grammatical systems are inherently variable or switching between systems must be taking place. These diametrically opposed positions as to the nature of variability have impeded consensus on how or even if variation should be formally represented according to modern theories of grammar. Theoretical Bilingualism, based on the assumption that grammatical systems can only feature categorical processes, is a concrete proposal for how the facts of variation can be explained using code-switching as explanation.

One advantage of Theoretical Bilingualism is that it resorts to a process with which we are familiar and which is widely attested, i.e. code-switching. However, a difficulty arises in that variation sometimes shows patterns that are not consistent with code-switching situations involving true bilingualism. Variationist study has frequently found that particular variants tend to occur more often in one grammatical context than in another. For example, Kroch (1989) showed that although the overall rate of change across contexts was constant for the rise of periphrastic do, it did occur in different grammatical contexts at different rates. There are several possible explanations why this should be true if the variation is due to code-switching. First, although we have yet to reach a consensus on what grammatical conditions on code-switching are, it may be that these grammatical conditions on variation represent the effect of grammatical constraints on code-switching. Without a clear idea of what these constraints are, however, this remains speculative. The other possibility is that speakers are choosing to code-switch only in certain grammatical environments for some other extralinguistic reason. To my knowledge, there is no evidence that speakers ever do this in cases of true bilingualism.

An alternative explanation to Theoretical Bilingualism is that optionality may arise from inherent variation which is systematically conditioned by both grammatical and extragrammatical constraints. Although it is true that code-switching exists both between different languages and between dialects of the same language, there are also clear indications that inherent variation exists as well (see Berdan, 1973; Labov, 1971). For example, in cases of shift between varieties, we usually find that features from both (or all) systems still occur in virtually all contexts and only differences in rate of occurrence serve to distinguish different varieties. Resorting to switching between grammatical subsets (or mini-grammars) is thus not necessarily the only principled explanation for systematic variation.

The question of whether variation can be explained as code-switching should be relatively easy to test empirically. For example, where an optional rule is incompatible with other features of the grammar, we should at least expect similar optionality in other areas consistent with a second system. In fact, this is what Kroch (1989) found to be the case with the rise of periphrastic do in English with variation also appearing in V-to-I raising. However, since Theoretical Bilingualism also allows switching between mini-grammars, evidence of a complete competing system is not always necessary to assert its existence. This is perhaps the most troubling aspect of the proposal, because when there is a situation where speakers can choose to switch between contradictory choices, one wonders how a researcher could determine which option represents the speaker's general linguistic system and which represents the mini-grammar. In other words, how can we know if a particular structure is a counter-example or evidence of another grammar?

Intuitively, we might consider that the option most consistent with patterns found for other grammatical structures in a language would be indicative of the general system. However, the situation is not necessarily that straightforward, as is revealed by a prominent instance of variability which underlies many recent minimalist proposals for English concord.

In modern standard English, variable subject-verb agreement in existential constructions like "There was three dogs in the yard" versus "There were three dogs in the yard" has been reported in various varieties of English both standard and non-standard (Eiskovits, 1991; Meechan and Foley, 1994; Tagliamonte, 1998). Given the feature-checking mechanism invoked to explain agreement, the choice without agreement would seem to be the one that is most consistent with the rest of the system, since the overt subject appears in post-verbal position. (i.e. not overtly in a position for feature-checking). That this analysis should hold for modern English is also supported by the variationist literature if we assume that simultaneous changes in related areas of the grammar can indicate the appearance of a single underlying structure or process.

A recent study of York non-standard English where concord is variable throughout the verb system found that although younger speakers show more concord in general and hence, more standard English patterns, their rate of concord in existentials is decreasing. Even so, recent minimalist analyses of English existentials, possibly influenced by prescriptive grammar, have chosen existential concord as indicative of the entire system and must invoke additional mechanisms in an attempt to fit it to the rest of the agreement system. Theoretical Bilingualism needs to specify what kind of evidence would be required to determine which variants are members of the same system. While Theoretical Bilingualism is an interesting proposal, it seems to me that in its present state, its explanatory power

is too strong to predict confidently all cases of variation and not demonstrably better than a theory assuming inherent variation conditioned by grammar.

References

Berdan, R. (1973). The necessity of variable rules. In R. W. Fasold & R. W. Shuy (eds.), *Analyzing variation in language*, pp. 11–25. Washington, DC: Georgetown University Press.

Eiskovits, E. (1991). Variation in subject-verb agreement in Inner Sydney English. In J. Cheshire (ed.), *English around the world*, pp. 235–255. Cambridge: Cambridge University Press.

Kroch, A. (1989). Reflexes of grammar in patterns of language change. *Language Variation and Change*, 1, 199–244.

Labov, W. (1971). The notion of "system" in creole languages. In D. Hymes (ed.), *Pidginization and creolization of languages*, pp. 447–472. Cambridge: Cambridge University Press.

Meechan, M., & Foley, M. (1994). On resolving disagreement: Linguistic theory and variation – There's bridges. *Language Variation and Change*, 6, 63–85.

Tagliamonte, S. (1998). Was/were variation across the generations: View from the city of York. *Language Variation and Change*, 10, 153–191.

The de-idealization of I-language

WAYNE O'NEIL
Massachusetts Institute of Technology, E39–212, Cambridge MA 02139–4307, USA.
E-mail: *Waoneil@mit.edu*

In his paper "Universal bilingualism" Roeper wants to account for the linguistic messiness of our minds by replacing the notion that the "state of the language faculty is some accidental product of varied experience, of no particular interest in itself, no more so than other collections of phenomena in the natural world" (Chomsky, 1995, p. 7) with a proliferation of I-languages: a grammar (perhaps only a mini-grammar) in the mind/brain to explain each class of contradictions that characterizes a person's unidealized knowledge of his or her language.

In this way, Roeper hopes to bring under I-language "what we call 'English', 'French', 'Spanish', and so on [that] even under idealizations to idiolects in homogeneous speech communities, reflect the Norman Conquest, proximity to Germanic areas, a Basque substratum, and other factors that cannot seriously be regarded as properties of the language faculty" (Chomsky, 1995, p. 11).

Roeper hopes in this way to explain a number of phenomena: the stages of language acquisition and of a language seemingly characterized by optionality, language change, exceptions, speech registers, etc. For example, an explanation of such archaic English words as "whereto" and "thereto" is said to follow from a grammar that allows the raising of pro-elements into Spec-PP, a fossil that remains from an earlier stage of the language – one that coexists in the mind together with a grammar that does not allow raising within prepositional phrases and that somehow dominates our language use.

Consider, now, two concrete examples that would seem to offer Roeper difficulty, the first one involving an exception to a general phenomenon in Japanese; the other involving a clear case of optionality in child language, with one of the options not evidenced in any known language.

First the exception: in Japanese there is a phonological process called *rendaku*, "sequential voicing", by which the initial segment of the second element of a compound is voiced:

(1) da + kashi → dagashi "cheap candies"
take + sao → takezao "bamboo pole"
maki + sushi → makizushi, for other types of sushi as well

It is well known that there are several conditions under which *rendaku* fails (Otsu, 1980; Uribe-Etxebarria, 1992). For example, *rendaku* is not possible in the second of each of the following pairs of compounds, where the second element is an ancient borrowing or a recent "foreign" one – either a Sino-Japanese (S-J) word (2b) or a later borrowing (3b):

(2) (a) ato + harai → atobarai "deferred payment" – /h/ being a reflex of an underlying /p/
(b) ato + kin → atokin "money left, balance" – kin being an S-J word

(3) (a) yasu + heya → yasubeya "cheap room"
(b) yasu + hoteru → yasuhoteru "cheap hotel" – hoteru, an English word

Another restriction on *rendaku* follows from the phonological prohibition expressed in Lyman's law, whereby in order for *rendaku* to go through, the second element of the compound may not contain any voiced obstruents:

(4) (a) oo + kata → oogata "big size"
(b) oo + kaze → ookaze "big wind"

(5) (a) oo + sakuranbo → oosakuranbo "big cherry"
(b) ha + kire → hagire "odd piece of cloth"

A third restriction requires that the compound be of the "right" type; i.e. that it not be a coordinate compound:

(6) (a) yama + kawa → yamagawa "mountain river"
(b) → yamakawa "mountains and rivers"

There are further complications, but (1) to (6) are enough to indicate to me where to draw the line between idealization and "factors that cannot seriously be regarded as properties of the language faculty" – between core and periphery: the constraint that depends on the bifurcation of the lexicon into native and non-native words cannot be part of the grammar; for Universal Grammar does not prepare one for this, or for irregularity in general. There is no grammar (mini or otherwise) that disallows sequential voicing in Japanese compounds that are of the "right" kind and whose second elements do not already contain a voiced obstruent simply because the second element is an unassimilated loanword. But Roeper appears committed to such a mini-grammar for Japanese, as he would apparently be for dealing with the division of English verbs into regular and irregular. However, it is a misinterpretation of the term "exception" to try to incorporate the periphery into the core in this way.

Roeper is also committed to a mini-grammatical account of what have heretofore been characterized as stages in language development involving optionality, the root infinitive construction being a striking case in point. Following Wexler we assume that "the alternation between finite and nonfinite forms of matrix verbs is characteristic of early child grammar" (Poeppel and Wexler, 1999, p. 5), and that children learning Dutch, French, German, English (perhaps) and non-pro-drop languages generally produce infinitives as the main verb of a root clause as well as well-formed finite verb root clauses. Thus in German we find both (7) and (8) (given in Poeppel and Wexler (1993, pp. 5–6), their (3a) and (4a)):

(7) I hab ein dossen Ball "I have a big ball"

(8) Thorsten Caesar haben "T. C. [= a doll] have (INF)"

Roeper's problem arises because he wants to argue that optionality entails the child's having two UG-consistent grammars, one of which s/he subsequently abandons. But in this well-attested case, there is no UG-consistent grammar of root infinitive constructions, and thus no mini-grammar to abandon.

For these and other reasons, I am skeptical of Roeper's theoretical bilingualism. The data on which it is built are frail. It constitutes, moreover, a move away from the sense "that much more substantial idealization is required if we hope to understand the properties of the language faculty" (Chomsky, 1995, p.7).

There is no doubt that we are all multilingual, given that our language environments have been varied and complex, but "it is hard to imagine that the properties of the language faculty – a real object of the natural world – are instantiated in any observed system" (Chomsky, 1995, p. 11).

References

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Otsu, Y. (1980). Some aspects of *rendaku* in Japanese and related problems. *MIT Working Papers in Linguistics*, 2, 207–227.

Poeppel, D., & Wexler, K. (1993). The Full Competence Hypothesis of clause structure. *Language*, 69 (1), 1–33.

Uribe-Etxebarria, M. (1992). A syntactic approach to morphophonological properties of compounds: On Japanese *rendaku* and (de)gemination phenomena in Dravidian languages. Manuscript, University of Connecticut.

Two grammars are better than one

CHARLES D. YANG
Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
E-mail: *charles@ai.mit.edu*

Roeper's theory of Universal Bilingualism (UB) consists of two central claims. The first claim is that children during language acquisition have simultaneous access to multiple grammars, i.e. principled syntactic variations made available by the innate knowledge of UG, and that these grammars compete to match the primary linguistic data. The second, and more specific, claim is that children's selection of grammars is governed by economy conditions suggested in the Minimalist Program (Chomsky, 1995). In this commentary, I will primarily focus on the first claim of UB and show, from a learnability perspective, its superiority over traditional acquisition models. I will also briefly discuss some developmental aspects of UB, in particular, the role of economy in language acquisition.

UB makes a radical departure from traditional models of language acquisition. Traditional models hold that the learner searches in the hypothesis space defined by UG to locate the target grammar, and at any time during this process, a *single* grammar is identified with the learner. A prototypical example is the triggering model (Gibson and Wexler, 1994), in which the learner moves from one grammar to another in a finite parametric space, driven by input sentences in the environment. Notice that the learner in the triggering model can radically change his current grammar on the basis of a single input sentence. We will see below that this all-or-none property of triggering results in a number of learnability problems, and that these problems do not arise for a UB learner.

The first problem with the triggering model concerns the existence of *local maxima* in the learning space. Local maxima are non-target grammars from which the learner is never able to reach the target grammar. Berwick and Niyogi (1996) formally characterize the pervasiveness of local maxima in Gibson and Wexler's 3-parameter space. More recently, Kohl (1999) shows, via computer implementation, that in a linguistically realistic, 12-parameter space, approximately half of the 4,096 grammars are not learnable. In the triggering model, the learnability condition simply is not met. A second and related problem has to do with the *ambiguity* of input sentences. Ambiguous input consists of sentences that are compatible with more than one grammar. (We abstract away from the obvious case of super-subset grammars that are inherently ambiguous, as they do not seem to arise in natural language.) As a result of ambiguous evidence, the triggering learner may be confused and select a wrong grammar. Although Fodor (1998) shows that there is unambiguous evidence for each of the eight grammars in Gibson and Wexler's 3-parameter space (and goes on to propose a revised triggering model that chooses grammars using only unambiguous evidence), it is doubtful that such an optimistic expectation holds for all natural language grammars (see Clark and Roberts, 1993). Lastly, the triggering model suffers from the presence of *noise* in the input, a situation that arises once a more realistic learning environment is taken into consideration. In particular, consider "noisy" input that is not compatible with the target grammar, but is compatible with some non-target grammar. For example, English-speakers, who in general use overt subjects, occasionally omit them in the so-called Diary drop constructions such as "seems good to me". This pattern, of course, is compatible with a Chinese/Italian-type grammar that does allow subject drop. Since the triggering model must associate the learner with a unique grammar, such input can lead the learner to switch back and forth between English and Chinese/Italian like a pendulum (Valian, 1990).

UB's multiple grammar approach to learning resolves the learnability problems that plague the triggering model. UB learning can be stated in the *variational* learning framework developed in Yang (1999a, 1999b). Each grammar in a UB learner is associated with a probability. When an input sentence is presented, the learner selects a grammar G with its associated probability P_G, and performs grammatical analysis (e.g. parsing). The success (failure) of G in analyzing the sentence increases (decreases) P_G. As Roeper points out, UB learning can be viewed as a Darwinian process in which grammars compete to match the input data. Formulated as such, UB learning is in fact one of the classic models of animal and human learning and decision making (Atkinson, Bower and Crothers, 1965).

One can quantify the "fitness" values of grammars and study the mathematical properties of the UB learner. One way to do this is to define the fitness of a grammar as the proportion of the input sentences with which it is compatible. Note that this is not to say that the learner keeps track of the frequency of input patterns; rather, she proceeds in the strictly Markovian fashion described above. The fitness (frequency) measures are extensional properties of grammars in a particular linguistic environment. It is possible to show that, in an idealized environment where all input sentences are generated by the target grammar T (see below for discussion of realistic and "noisy" environments), the UB learner converges to the target unique grammar T (i.e. $P_T = 1$) with probability 1 (Narendra and Thathachar, 1989; Yang, 1999b). That is, there is no local maxima in UB learning, and convergence is guaranteed.

To see how UB learning circumvents the ambiguity problem, consider the following population of grammars in a German environment:

(1) (a) German: SVO, OVS, XVSO
(b) English: SVO, XSVO
(c) Hixkaryana: OVS, XVSO
(d) Irish: VSO, XVSO

It is clear that none of the patterns in (1a) can alone identify the German V2 grammar unambiguously, as each of the three non-target grammars in (1) are compatible with *some* German input sentences. However, each of them is also *incompatible* with a positive amount of input, and will be punished as a result. The target grammar, being 100 per cent compatible with all the input, eventually emerges to eliminate all the competitors. Therefore, the learnability condition is met in UB learning despite the lack of unambiguous evidence; this is a stronger result than Fodor's (1998) revised triggering model, which requires unambiguous evidence for convergence.

The problem induced by noisy input receives a new interpretation in UB learning. First, as Roeper points out, one must recognize the inherent nature of "noise" in natural language: any two properties present in a language that are compatible with a single grammar can be viewed as a special form of noise from a learnability prospective. In a classic paper, Weinreich, Labov and Herzog (1968) argue that it is unrealistic to study language as a "homogeneous object", and the "nativelike command of heterogeneous structures is not a matter of multidialectalism or 'mere' performance but is a part of unilingual linguistic competence" (p. 101). This implies that the primary linguistic data children are exposed to are inherently contradictory ("noisy").

The behavior of UB learning in such a realistic and linguistically non-inform environment can also be formally studied. One can show (see Yang, 1999b) that when neither of two competing grammars is 100 per cent compatible with the input, the UB learner converges to a stable combination of probabilities expressed in terms of their fitness measures (i.e. the percentages of the input sentences with which the two grammars are respectively compatible). A number of predictions from this are confirmed by empirical studies, including, as Roeper notes, Kroch and his colleagues' observation that speakers during language change exercise two distinct grammars. Note that contradictory data lead to such desirable consequences only in UB learning, which intrinsically entertains multiple grammars; they lead to serious learnability problems in the one-grammar-only triggering model, as noted earlier.

Following my enthusiastic support for Roeper's UB (I also refer the reader to my own work cited above for additional developmental evidence showing the reality of multiple grammars), I am skeptical of his second claim that learners use economy conditions to select among competing grammars: simpler grammars are preferred. From a learnability perspective, economy conditions are not necessary for learning to be successful. The grammar that is most compatible with input data emerges as the winner, regardless of its complexity. Empirically, the use of economy leads Roeper to a stage-by-stage view of language development: starting with a Minimal Default Grammar (MDG), more complex grammars gradually become accessible to the learner. Consider Roeper's example on the acquisition of the Case system:

(2) (a) MDG: me want.
(b) MDG + target: me want and I want.
(c) target alone: I want.

(2) does not quite accurately characterize the development of the Case system. Children do not start out by using "me want" exclusively, contrary to what (2a) suggests. For example, Schütze (1997, p. 225) reports that in an English child Nina's earliest speech (1;11–2;5), only 13 per cent (137/1040) of 1sg pronoun uses are errors involving "me" and "my" (although it is possible that when the relevant recording started, Nina had already passed the MDG stage). It seems plausible to assume that both grammars are available from the outset, and, following Roeper's suggestion, the learner gradually eliminates MDG on the cumulative evidence of agreement patterns in the target language. Therefore, the mastery of the Case system is preceded by a mixed use of both MDG and the target grammar, rather than by a stage in which only MDG is present.

Roeper's UB theory introduces an important concept in the study of language acquisition. The assumption of multiple grammars yields some desirable learnability results, an improvement over the standard triggering models. Furthermore, as is clear in Roeper's account for child Case errors and many other examples presented in his paper, multiple grammars promise to offer deeper explanations for the non-uniformity in child language data that had to be attributed to "optionality" or "performance" in models that identify the learner with a single grammar. Much of the previous studies in theoretical language acquisition ought to be reformulated and reanalyzed if UB proves successful.

References

Atkinson, R., Bower, G., & Crothers, E. (1965). *An introduction to mathematical learning theory*. New York: Wiley.

Berwick, R., & Niyogi, P. (1996). Learning from triggers. *Linguistic Inquiry*, 27, 605–622.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

Clark, R., & Roberts, I. (1993). A computational model of language learnability and language change. *Linguistic Inquiry*, 24, 299–345.

Fodor, J. D. (1998). Unambiguous triggers. *Linguistic Inquiry*, 29, 1–36.

Gibson, E., & Wexler, K. (1994). Triggers. *Linguistic Inquiry*, 25, 407–454.

Kohl, K. (1999). An analysis of parameter learning in linguistic spaces. S.M. thesis, MIT.

Narendra, K., & Thathachar, M. (1989). *Learning automata*. Englewood Cliffs, NJ: Prentice Hall.

Schütze, C. (1997). Infl in child and adult language: Agreement, case and licensing. Ph.D. thesis, MIT.

Valian, V. (1990). Null subject: A problem for parameter-setting models of language acquisition. *Cognition*, 35, 105–122.

Weinreich, U., Labov, W., & Herzog, M. (1968). Empirical foundations for a theory of language change. In W. Lehmann (ed.), *Directions for historical linguistics: A symposium*. Austin, TX: University of Texas Press.

Yang, C. (1999a). Formalizing language development. In *The proceedings of the 35th regional meeting of the Chicago Linguistic Society*. Chicago, IL.

Yang, C. (1999b). A selectionist theory of language acquisition. In *The 37th annual meeting of the Association for Computational Linguistics*. Maryland, MD.

 Bilingualism: Language and Cognition 2 (3), 1999, 204–208

AUTHOR'S RESPONSE

Redrawing the boundaries of grammar*

THOMAS ROEPER
Department of Linguistics, University of Massachusetts, Amherst, Mass. 01003
E-mail: *roeper@linguist.umass.edu*

Restatement

The concept of Theoretical Bilingualism turns the problem of the boundaries of grammar upsidedown. Instead of characterizing the "grammar" of a speaker as whatever system lies behind a language, a minimalist approach (as I see it) argues that the real principles of grammar are so abstract and narrow that no single non-contradictory representation lies behind a "grammar" of a real language. So, although principles remain, when different parametric choices are linked to different lexical items (in a Numeration) then contradictory choices are inevitably made. In effect then, the heterogeneity of a single "grammar" is like the heterogeneity in being bilingual.

Is there then a well-defined object of inquiry? This is a question that many commentators asked and was a central query in the recent discussion by Jürgen Meisel at GALA (1999).

If we start from a biological perspective, then anything that enables rapid integration of information is the natural object of inquiry, whatever we choose to label it. The ultimate goal is to understand the mind by characterizing the neurology that can do rapid information integration. It will be captured in a mathematically defined formula, just as physics uses formulas. If a human being can do code-switching in defined places in a sentence (Woolford, 1983), then a description of a human ability in neurological terms is necessary to capture it. If it involves two different grammars, then there must be a stateable relation between those grammars.

If there turns out to be no recognizeable core notion of grammar, it is far from a poor result. In fact, one might say that the current effort to define a minimal core grammar and its interaction with interfaces is a move toward the dissolution of the traditional notion of a well-defined grammar. Scientific progress has been persistently marked by the jettisoning of apparently crucial concepts in favor of more refined ones. There can be no doubt that we have obtained deep insight into grammatical principles, such as bound variables and long-distance movement, whatever the boundaries and interfaces eventually look like. The situation is analogous to the view Chomsky (1982) has suggested, that the concept of "human body" will prove useless, particularly as we focus on the disparate principles that lie behind various organs and their connections to mind.

* Thanks to Kyle Johnson for discussion, and to the commentators themselves for their stimulation.

UG evolves

In a sense, whenever a new proposal about UG is offered, the range of possible forms of bilingualism change. Chomsky (1967) has argued that every notational proposal is a claim about how the human mind works. If we allow parentheses to indicate optionality, then optionality can be captured in a single grammar. If we rule it out, just as Chomsky and Halle (1968) expressed doubts about the powerful mechanism of angled brackets, or Chomsky (class lectures) and Reinhart (pc) have challenged the use of indices recently to represent binding, the set of possible grammars, hence possible forms of bilingualism, change. That said, it remains the goal of minimalism to reduce the central options of grammar – consigning some systematic phenomena to the interface level – and therefore the bilingual perspective becomes ever more relevant, even though precise predictions change as UG changes.

Is it possible that a grammar could be defined as something that could be instantly created? If a person, as Peter Coopmans suggested, adopted a French accent for an English sentence, then he would generate, momentarily, a different grammar. I think that is correct. But it does not assert that variation is unlimited (see Henry, 1997 for relevant discussion). I think it would be impossible for a person, for every three words, to adopt a French, then English, then German accent in fluid speech. Likewise it is not possible to do code-switching for every other word in a sentence, although it is possible to switch grammars in a sentence midway. (Since it is a mathematically easy concept, it would be easy to program a speaking computer to shift accents serially for each sequence of three words, which humans could not do at a normal rate of speech.) So serial three-way accent variation would not be a possible form of bilingualism. This traditional reasoning already shows that there are upper limits to the heterogeneity that human beings can incorporate in on-line use of multiple grammars.

V-2 again

Take another example: suppose we assert that one at any given moment selects a particular configuration from a set of possible grammar ingredients to form an instant grammar. The possibility of on-line bilingualism would then be circumscribed by the notion of what can be instantaneously combined. What could be excluded? One cannot do residual V-2 and V-2 at the same time (2a,b). To pursue our example, one can say:

(1) (a) "No" says John
(b) Could John say "No"
(c) "No" could John say

but not:

(2) (a) *"No" could say John
(b) *Could "no" say John

This could constitute an excluded form of on-line bilingualism, even for a speaker who has both residual V-2 (move aux) and V-2 (move main verb); see Tracy (1996) for more examples.

These cases show (i) that we have not abandoned all constraints, and (ii) that we can use traditional methods to approach bilingualism and refute the view that the problem is hopelessly obscure (e.g. O'Neil's apparent view, which I think promotes a deleterious pessimism, that no meaningful generalizations about bilingualism are currently possible). The general question: where does systematicity lie in bilingual ability? remains the challenge. It is of a piece with the challenge of "constraining" over powerful grammars that has been a traditional goal. I have tried to frame the question, but I certainly have offered only hints of the answer.

An important focus should remain on the contrast between overt and covert movement and the claim that Logical Form is universal. This hypothesis would explain how an able English speaker, whose first language is Dutch, reports that she was confused by a sign on a train that says "all the doors will not open" which in English means "some", but translated into Dutch means "all", because English allows covert movement of negation generating "not all" from "all not". Such a confusion is only possible if the Dutch LF is dominating the English one.

Commentaries

The commentaries in general advance the discussion. Hawkins emphasizes the view, which I share, that all of the distinctions that can lead to on-line bilingualism can be captured in the lexicon (with more than one representation of C). Still, the consequences would be carried out in the computational part of the grammar. So both components would be involved.

Hawkins supports the next logical step in the theory by emphasizing that UG would be conceptually cleaner if one can show that not only are all nodes linked to semantic concepts, but that all movement operations have a semantic effect. If we look at more subtle semantic effects and link them to formal features in the syntactic system, this claim may be sustainable.

Hawkins makes just the connections I was hoping would emerge: he observes that the combined notions of a Default Grammar and bilingualism directly predict the results obtained by White on L2-speakers preferring do-insertion to verb-raising.

Many commentators (Hulk, Haider, O'Neil) fixed on eliminating "optionality" as an important prediction. In retrospect, it seems more like a terminological one. There is no question that "optionality" is present as a descriptive phenomenon. The question is only one of whether we choose a formal mechanism, like parentheses, to claim that that it is a part of our mental analytic phenomena, or whether we avoid it, allowing each option a lexical representation. To give the matter an empirical dimension: do we have (a) one grammar with parentheses, or (b) mini-bilingualism (two subcategorizations = two verbs):

(3) (a) give (NP) NP (PP)
or two forms:
(b) give NP
(c) give NP NP

Close research (Randall, 1992) suggests that the latter form, two forms rather than optionality through parentheses, is consistent with subtler approaches to meaning, and captures the fact that certain forms require both NPs, as only (b) allows (Oehrle (pc)):

(4) (a) *John gave a headache (to me)
(b) John gave me a headache

The parentheses around the PP are also captureable through "redundancy rules" which would invoke a different notation.

Root infinitives are another case. Current research has shown that root infinitives are not optional forms of tensed infinitives. The latter seem to occur only where events are present (e.g. Wijnen, 1997). Each instance, then, of putative optionality, will call for a separate analysis, but the direction of the claim should be clear.

Meechan's discussion presents "inherent" variation as an alternative. It is not clear to me what the term implies at a more technical level. It would seem to reduce to lexical variation, which is in fact what we are advocating. The example of variation in there-insertion is a nice case to consider. We have to examine the phenomenon with greater subtlety.

In a paper (Schafer and Roeper, 1999) we argue that at the earliest stage there is a form of presentational expletive that has both the features +loc and +expletive:

(5) (pointing) there's two dogs

One possibility is that the agreement with the associate (rightward subject) is triggered when pure expletives arise [+exp], as in negative contexts:

(6) there are no hats

At this point there-expletive and there-locative would be split. When it is split, the locative "there" would be seen as contentful and call for singular, while the expletive would trigger agreement with the associate. (See Hollebrandse, 1997; Moore, in preparation, for more evidence in behalf of the notion that splitting is an acquisition operation.)

It remains possible that further meaning differences exist. In general the singular "is" gives preference to the collective reading of plural nouns and the plural "are" favors a distributed reading, although pragmatically both are available.

New examination of fundamental assumptions is always

welcome. Haider presents what seems like a radical departure in his input guided system. However it seems to me to be just a revived version of the intuition behind behaviorism that the organism is molded by stimuli. Chomsky's insight that the organism selects a tiny part of the possible stimuli and gives them a tiny range of possible interpretations remains, I think, essentially unchallengeable. The "poverty of stimulus" argument, namely that the input underdetermines the child's conclusions, must be clearly overturned for his perspective to have power. Nothing has overturned the idea that the child with UG guides the process, though common sense says it is the other way around.

In any case, even within Haider's system, a much more critical question must be addressed: children are exposed to highly ambiguous data which could easily mislead them. It is quite unclear to me that his system can cope with the ambiguity in the input provided to the child.

It is surprising that in his detailed commentary, he makes no direct use of the theory he advocates, giving instead standard analyses of the phenomena.

First the argument for optionality refers to the inversion of auxiliary and "have". Again I suspect a meaning difference, suggesting no optionality. Consider these cases:

(7) (a) John insisted that it not have been eaten before he left
(b) ?John insisted that it have not been eaten before he left

While (a) seems to be a command, (b) seems to be questionable, and if acceptable, the description of a conviction. The distinction is hard to grasp but feels just like the ones discussed with "already have" and "have already" where temporal anchoring differs. Whatever the answer, the conclusion that we have a clear case of optionality is unwarranted.

The statement "do-insertion is a reflection of a single, exceptionless property of English: main verbs do not move to functional head positions ... features may be lowered" is an older view which is certainly not true if "lowering" is eliminated, as many argue. Haider offers no argument to explain the acquisition data where do-insertion is overgeneralized.

Finally it may be that the core/periphery distinction continues to be valuable, but this is not necessarily the case. Chomsky, for instance, has essentially disavowed it in advocating a more intricate connection to interfaces.

Ayoun's interesting extensions raise a very deep question, which in a sense, could undermine many of my specific claims, though not, I think, the idea. As our claims about UG change, what is formulable within a single rule changes. The arguments on behalf of TB, as with those advanced by Ayoun, could fall apart if the theory changed in a way that recaptured two options under a single system. This happens easily if a new feature is identified which motivates movement. For instance, suppose reflexives raise but not pronouns (as has been suggested in various places), then they might be responding to two versions of a verb, one reflexive-marked, the other not, as Reinhart and Reuland (1993) have suggested. Now this reflexive marking would be represented in the lexicon and might in fact involve different meanings. It has been pointed out by Kuno (pc) that there is a slight difference in meaning between:

(8) (a) John put a blanket under him
(b) John put a blanket under himself

He claims that in the (8b) case the blanket is closer to him physically than in the (8a) case. One might somehow shift this over into a pragmatic component, but under the proposal mentioned above, it would be in the grammar and we would have two different meanings for these expressions, which would lead us to keep them in the same grammar.

However, in the formulations offered by Ayoun we find indeed that there is clear evidence for two grammars rather than a looser single one. The argument on behalf of TB from a minimalist perspective is that where one seeks a narrower conception of grammar, the likelihood that multiple grammars are necessary increases. The essay seeks to articulate that possibility as one to be kept in mind in the effort to minimize what a possible grammar can be.

Yang is pursuing the logic of UB at the level of mathematical learnability. His results should shed light on whether there are mathematically stateable limits on the interaction of notational choices for a range of different assumptions about what representational powers the mind provides. I would encourage this line of research to focus more on the mathematical implications of explicit minimalist or non-minimalist choices. How does minimalism differ in its implications from various choices in Categorial Grammar (to which minimalism is closely related) (Chomsky, personal communication).

Yang's mathematical approach to the ambiguity of input is promising, particularly if it can accommodate the following differential feature of "local maxima" which was emphasized by Weissenborn and Roeper (1990). The suggestion there was that subparameters arise in response to ambiguous input, allowing opposite parametric choices. (This predicts, on the basis of some minimal evidence, that "the emergence of the unmarked", as argued in optimality theory, would occur.) But the two parametric choices were differentiated: the primary parameter becomes productive, while the minor parameter choice is linked exclusively to the lexicon. This captures the fact that empty subjects occur only with specified verbs in English (seem, look like) while it is productive in Italian. Consider the statement "... not the case that the child keeps track of the frequency of input patterns. The fitness (frequency) measures are extensional properties of the grammar." Now, can we include precisely the notion of a lexical link in these "extensional properties"?

Though unacquainted with modern work on animal learning, it seems unlikely that animals would have more abstract learning systems. One would expect instead that they would have more hard-wired task-specific learning capacities. Until shown otherwise, I would expect that is true. Now suppose we can show that a dog will chew a

bone and a plastic object that looks like a bone. Has the dog really generalized? Or are they hard-wired to chew bones, but simply can be fooled, rather than the idea that they "learned" to include hard plastic in an abstract definition of chewable objects?

Should "local maxima" do all the work or is there a useful role for minimalist concepts of simplicity? We must still explain those instances where grammars are abandoned in the process of acquisition. If we understood the matter at the right level of detail, it could reveal that a child silently considers and rejects 20 grammars a day. We still have very little grasp of the micro-structure of the mental process of grammar construction, but it would be surprising if economy played no role in them.

Hulk suggests that "optionality" must be divided into two kinds: developmental and stable. Why should certain forms of optionality disappear in hours during acquisition while others seem to hang on for 500 years? This descriptive division suggests that our account – leaning on lexical representation – may not capture all of the variation at work. This is a good challenge, to which I have no ready response, other than to urge that we anchor our claims in specific examples.

What remains a crucial question, which might in fact be captureable in terms of conditions on local maxima, is the trigger whereby something acquires a categorial rather than lexical definition. If verbs of quotation, "be", "have", and a few others are subject to V-2, then what triggers the shift to a definition in terms of V? Perhaps it is not increasing numbers, but diversity of types. Perhaps when the grammar of the child involves verb categories of totally different types (verbs of motion and of speaking), then the rule becomes general. Another possibility (from a non-specialist in historical linguistics) would be that it is the encroachment of residual V-2 that turns generalized V-2 into a lexically limited phenomenon.

O'Neil considers the presence of "wherefrom" to be just accidental. But now we must examine what that term means. In a sense, the invention of any word in a language is just accidental. We don't have to have a word such as "convertible" or even a modal such as "darf" (permit) in English (the proof of which is that only German has the modal "darf").

What is significant is not some degree of "accidenticality" but only whether we can locate systematicity in interaction among grammars. Note that under the account of van Riemsdijk (1978), in which he argues that Prep has a Spec in German, it is not surprising if both clauses and PPs allow a complementizer. If they do, then they should function as a barrier when filled. A contrast arises, though subtle, between (a) and (b):

(9) (a) ?When did John say that to home he had gone t
(b) *when did John say whereto he had gone t

It is explained if we assume that the wh-word in "whereto" occupies a Spec position that blocks extraction in (b). If it were a pure relic, with no internal analysis left, then we would expect no interaction with the grammar and (a) and (b) would be the same. Presumably the whole PP (whereto) moves into Spec of CP in order to satisfy checking. The necessity of this is indicated by the fact that the wh-subcategorized in Spec of CP by "wonder" can be satisfied:

(9) (c) I wondered whereto you had gone
(d) *I wondered a picture of whom you saw
(e) I wondered who you saw a picture of

The contrast between (9d) and (9e) indicates that "wonder" immediately dominates wh- without Seeking Sublabel, as we argued in UB.

Still lexical restrictions clearly apply. One cannot prepose across any preposition:

(10) *did he say where-between he travelled.

This calls for an interaction between the grammar of wh-movement and the limited grammar of wh- in Spec of P. Unless one assumes that all categories have Specs as barriers, which is reasonable, but not implied under minimalism (where positions are constructed by movement), then one must arrive essentially at an interaction among mini-grammars, which equals bilingualism.

Ellipsis, sluicing and shortest move

Let us further explore the nascent discipline of TB as a branch of UG studies by looking at the grammar of preposing inside PPs a little more closely, despite its limited productivity. Consider these sentences:

(11) (a) I gave the money to someone, but I forgot who to?

This could be analyzed as an interesting form of sluicing rather than simply VP-ellipsis:

(12) (b) who (gave the money) to (trace)

with intricate LF reconstruction. (See Johnson, 1997 for a careful study of gapping and ellipsis.) But we find here that the same restrictions seem to apply:

(13) (c) *I found him among old people, but I don't know exactly who among.

If this is a case of sluicing where all the missing material is (somehow) reconstructed, then we should reconstruct a sentence like (14a) for (14b):

(14) (a) who did you find him among *trace*
(b) I found him among old people, but I don't know exactly who I found
him among *trace*

But (13c) is worse than (14a) or (14b). Suppose we argue that the derivation is different. It follows these several steps. First we have movement of the PP:

(15a) I gave it to someone, but I don't know to whom (I gave it trace)

Now sluicing occurs, deleting (I gave it trace), and then preposing within the PP occurs:

(15b) but I don't know who to.

Now we have the perspective from which to argue that

(16) *... know exactly who among

is ungrammatical, because it does indeed require movement to Spec inside the PP, which is apparently unavailable for complex forms like "among". How do we capture its unavailability? One possibility is to argue that the prefix "a-" is actually a kind of Spec, as it seems to be a modifier adding aspectual information in other cases which block certain syntactic possibilities like "*the asleep boy/*the afraid boy". Thus the filled Spec would block wh-movement.

A problem remains, however, in this analysis. How is it possible to have:

(17) what did you find the boy among

if (18=13c) is out:

(18) *I found the boy among some things, but
I don't know exactly who-among.

Under minimalist restrictions, one would move through the Spec-PP in order to get to Spec-CP, which a P like "among", with filled Spec, would block. If, however, we distinguish between Shortest Move and Shortest Attract, as proposed by Richards (1997), then we can argue that (17) is the result of Attraction and not just Move. Therefore, in effect, it does not have to follow the same path through the structure as Shortest Move.

Now let us ask the TB question: would we predict that there is an interaction, as with other LF phenomena, across grammars? German has more forms like ("womit, wonach") than English, but it does not have VP-ellipsis, although it does have sluicing. Another window of analysis for UG is opened if we ask how easily such structures can be transferred from one grammar to another. A more detailed analysis is called for, but the outlines of the approach should be clear. Much of the literature has analyzed where transference occurs in L2. Here we are trying to make theoretically motivated predictions instead. Moreover, one can imagine experimentation (perhaps adapted from the work on the acquisition of wh-movement that I have pursued with Jill de Villiers, 1990, 1995) that could address such questions.

Conclusion

The commentaries take up the challenge of articulating the more precise interactions among grammars which will be the future of TB. The primary point I seek to emphasize is that the deliberate narrowness of operations permitted in minimalism, makes the articulation of multiple grammars more important as a theoretical goal. Problems like "is optionality real" are really technical questions about what representational devices grammars will allow. One hopes that the actuality of bilingualism could bring very pertinent evidence to bear upon such decisions. The essay simply shows that the learnability advantage obtained by eliminating optionality dovetails with the arguments for bilingualism.

What is the status of bilingual data? The reality is that theorists are overwhelmed with diverse data coming from the languages and dialects of the world. This is no different from biology or physics which are both overwhelmed with unexplained phenomena. It is difficult to know what belongs in the forefront. Careful examination of binding and movement within English has led to profound insights. When case-marking became a focus of attention, other languages moved to center stage. Ellipsis is a domain where TB perspectives appear to be particularly relevant. Perhaps as the concept of *interfaces* becomes critical, though it is intended to deal in part with modules within grammar, it will naturally place a new and deserved focus on TB and the challenge of representing the superstructure that carries multilingual knowledge.

References

Chomsky, N. (1967). Some general properties of phonological rules. *Language*, 43, 102–128.

Chomsky, N. (1982). *The Generative Enterprise: A discussion with R. Huybregts and H. van Riemsdijk*. Dordrecht: Foris Publications.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper and Row.

de Villiers, J. G., Roeper,T., & Vainikka, A. (1990). The acquisition of long distance rules. In L. Frazier & J. G. de Villiers (eds.), *Language processing and acquisition*, pp. 257–297. Dordrecht: Kluwer.

de Villiers, J., & Roeper, T. (1995). Barriers, binding, and the acquisition of the NP/DP distinction. *Language Acquisition,* 4, 73–105.

Henry, A. (1997). Dialect variation, optionality, and the learnability guarantee. In A. Sorace, C. Heycock & R. Shillcock (eds.), *Gala Conference on Language Acquisition 1997*, pp. 62–68.

Hollebrandse, B. (1997). The split C projection: Evidence from the acquisition of sequence of Tense. In A. Sorace, C. Heycock & R. Shillcock (eds.), *Gala Conference on Language Acquisition 1997,* pp. 73–79.

Johnson, K. (1997). What VP ellipsis can do, what it can't, but not why. Manuscript, University of Massachusetts.

Moore, D. (in preparation). Dissertation on the acquisition of comparatives. University of Massachusetts.

Meisel, J. M. (1999). On the possibilities of becoming a monolingual but competent speaker. Paper presented at GALA 1999.

Randall, J. (1992). The Catapult Hypothesis, an approach to unlearning. In H. Goodluck, J. Weissenborn & T. Roeper (eds.), *Theoretical issues in language acquisition*, pp. 93–138. Hillsdale, NJ: Erlbaum.

Reinhart, T., & Reuland, E. (1993). Reflexivity. *Linguistic Inquiry*, 24, 657–720.

Richards, N. (1997). What moves where, when, and in which languages. Dissertation, MIT.

Schafer, R., & Roeper, T. (1999). The role of the expletive in the acquisition of discourse anaphora. Manuscript. University of Massachusetts.

Tracy, R. (1996). *Sprachliche Strukturentwicklung*. Tübingen: Gunter Narr Verlag.

van Riemsdijk, H. (1978). *A case study in syntactic markedness*. Dordrecht: Foris Publications.

Weissenborn, J., & Roeper, T. (1990). How to make parameters work: Comments on Valian. In L. Frazier & J. de Villiers (eds.), *Language processing and language acquisition*, pp. 147–162. Dordrecht: Kluwer.

Wijnen, F. (1997). Incremental acquisition of phrase structure: A longitudinal analysis of verb placement in Dutch Child Language. Manuscript, University of Groningen.

Woolford, E. (1983). Bilingual code-switching and syntactic theory. *Linguistic Inquiry*, 14, 520–536.