Universal Bilingualism

Thomas Roeper

Department of Linguistics

University of Massachusetts

Amherst, Mass. 01003

[roeper@linguist.umass.edu]

May 1999

Abstract: 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.

1.0 Introduction[1]

       We argue that a narrow kind of bilingualism exists within every language.[2]   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.[3] 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.

---

[2]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.
[3]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.

Much of what we shall claim about multiple grammars has been claimed before. Two features distinguish our approach from previous ones: 1) we use the concept of Theoretical Bilingualism to capture recalcitrant features of first language acquisition, in particular, optionality and lexical variation and 2) we utilize Minimalist theory to state in terms of economy where bilingualism within a language is predictable .[4]

The details of bilingual variation often receive an accurate description 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. In traditional terminology, both options of a mutually exclusive parameter are chosen.

---

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

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.[5] The logic of learnability theory is this 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, which 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.

---

[5]See Berwick (1985)

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 <u>a grammar</u> 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 will make is that where two grammars are present, one may represent a <u>Minimal Default Grammar</u> definable in terms of economy.[6] Nonetheless, most of the questions about when exceptions survive or disappear remain.

1.1 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.

1.2 An Example

---

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

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").[7] There are several logical approaches to this phenomenon.

1) Each form ("I want" and "me want") represent 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).[8]

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. [9]

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

---

[7]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.

[8]See Roeper and deVilliers (1992), Abdul-karim and Roeper (1996), and Schutze (1997) for discussion and references for this phenomenon.

[9]Powers (1996) argues that forms like "I want" precede and co-exist 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 like [IP I_i [VP pro$_i$ $_{[want]}$]] with a chain between the two subjects. Any analysis must, however, explain why these structures should co-exist. 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.

Roeper and Rohrbacher (to appear), 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 for 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.[10] First we will discuss the role of inference in the use of incomplete grammars.

1.3 Interface Economy: Limiting the role of Inference

---

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

Adults, like children, are more or less explicit depending upon the social occasion. For instance, if one enters a store and says either a) "milk" or b) "I want milk," both utterances have the same ultimate meaning but (b) is clearly more explicit. Situational inference, not written into the grammar, makes (a) 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."

1.4 Economy of Representation

It is important to recognize that no regular input justifies the expression "me want", or G2.[11] It is effectively a spontaneous expression derived from innate

_____

[11]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.

knowledge of Universal Grammar.[12] 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-karim (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: 1) economy of representation and (2) 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).

1.5 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.

_____

[12]Note that Bickerton (1981) also claims such structures for Creole languages.

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.

1.5.1 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 nounphrase 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 <u>she sings</u> 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? This is a very deep question to which there is no straightforward answer. 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.

1.6 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.[13] 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.[14] If Agreement is optional, then hearing an example like "he walks" cannot establish that it is obligatory.[15]

1.7 The Link to Social Registers

"Pro-drop" languages allow Null subjects ("goes" instead of "he goes") and they are commonly differentiated 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").[16] The missing subject is either a special rule, called "Diary Drop" (Haegeman (1993)), or it is the

---

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

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

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. [17]   What is of interest is a) that the choice of grammar can be linked to social register, and b) 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

---

[17]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.

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 co-exist 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.

1.8 Theoretical Sketch

We provide here a perspective on the relations between principles of economy, a Default Grammar, and a Particular grammar. [18] This is then the formal source of one form of bilingualism:

a. Universal Grammar defines a set of Default representations

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

---

[18]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 deVilliers and Roeper (1992) for use of the notion of Default case, and more recently Schütze and Wexler (1996).

than elaborated particular grammars.

c. The Particular Grammars and the MDG grammar may or may

not be incompatible.

d. Different grammars can be <u>localized</u>:

1) in lexical classes

2) 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) NP's below DP's 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.[19] If a grammar lacks Agreement, then it is a direct reflection of MDG.

2.0 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":

8) A single salad does not a dinner make.

_____

[19]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 Schutze and Wexler (1996) who extend the argument for the optionality of Agreement.

This form is generalizeable:

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 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.[20]

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?

2.1 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

---

[20]A current theory by R. 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 form 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.

child's gradual exposure to the language, this is a logically significant possibility. We find that both <u>be</u> and <u>have</u> invert:

      17) a. is he here

          b. have you a dollar[21]

In sheer frequency terms, the child hears a significant portion of V-2 expressions (like "what is that?").[22] In 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 (<u>what is that</u>) and non-V-2 (<u>what did he say</u> not *<u>what said he</u>). 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 <u>say</u> and <u>shriek</u> 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.

---

[21]This form is becoming fairly rare in modern American English, but less so in British English.
[22]See Takahashi (1990) and Stromswold (1995) for arguments that inversion must be present in these cases.   Note that demonstratives cannot function as predicates: *<u>a fish is that</u>. Therefore <u>what is that</u> must come from <u>that is wh-something</u>.

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 <u>have</u> and <u>be</u> 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 <u>be</u>: [23]

20)   "do it be colored"

        "you don't be quiet."

        "Allison didn't be mad"

        "this didn't be colored"

        "did there be some"

---

[23]See Roeper (1993) and Davis (1987) and references cited therein for sources. Moreover, adults in American English today are progressively avoiding inversion with <u>have</u> , preferring (i):

i) Do you have a dollar?

We are in the midst of a form of language change with respect to the verb <u>have,</u> which is 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 <u>have</u> 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 <u>have</u> is strictly limited lexically.

```
"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[24] 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:

---

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

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 in Anglo-Saxon, Latin, and Greek origin. [25] 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 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.

---

[25]For instance, see Randall (1981), for a discussion of affixation.   She shows that speakers know that <u>civility</u> is possible but *<u>evility</u> is not since the latter is Anglo-Saxon and not Latinate. However the Anglo-Saxon affix -<u>ness</u> can appear with both forms: <u>civilness</u> and <u>evilness</u>. How did <u>-ness</u> lose its Anglo-Saxon moorings and become productive for all nouns?

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.

2.2   Language Interference

Is there an abstract answer to this question: How can grammars interfere with each other? Code-switching and lexical borrowing is evidence of where grammars can connect and interpenetrate.   But we do not know, hand, 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 occur 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: 1) the operation occurs overtly in German, but not in English, and 2) 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. Chomksy (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.[26] 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[27]:

    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) Hanns 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

---

[26]See Giorgi and Fabiesi (1997)

overtly, then finiteness is not fixed. [27]   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

nounphrases.   We now argue that visible movement gives definite reference, via

temporal anchoring, to verbphrases.

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 came in the room with a new toy. John says "awesome" over and

over.

---

[27]Meisel (1994) represents Tense as distinct from Finiteness, locating
Finiteness in C, following Platzack and Holmberg (1989), and Hakannson (1998)
argues that children fail to represent Finiteness as opposed to Tense.
Moreover, Hirschensohn (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.

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.[28]   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 <u>have</u> has raised above the adverb <u>already</u>, 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 <u>have</u> 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.[29]

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 (a) overuse of the progressive, or (b) 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

---

[28]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.

[29]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.

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.

3.0   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.[30]   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 Nounphrases, 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.[31]

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]The economy of representation approach is pursued in work by Roeper (1996) and Roeper and Rohrbacher (1994) and Rizzi (1995), who formulates the idea as "Avoid Structure".
[31]See deVilliers and Roeper (1995) for discussion.

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.[32] Therefore, as we argued above, sentences of the form <u>me want</u> arise among a number of English-speaking children when they recognize <u>me </u>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.

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 (1991)) 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$_i$" [33]

---

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

[33]Note that the view that this is purely a processing effect would not explain sensitivity to quantification. Resumptives are much worse for quantification:
   i.*No book that when I read it I was completely confused.

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.

4.0 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

---

(Demirdache (1991))

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 <u>do</u>-insertion is just the Spellout form of a Strong affix. In other words, the form <u>did</u> is just the way we pronounce <u>-ed</u> by itself (following a suggestion by H. Lasnik (pc)). Under this hypothesis, however, English contains both Strong independent affixes linked to <u>do</u> 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 <u>do</u>-insertion form is in fact preferable.[34] 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

---

[34]See also Caviar and Wilder (1996) for similar arguments applied to Serbo-Croatian.

argument is this: the tense marker in <u>talked</u> is buried in the verb, while the tense marker in <u>did talk</u> is explicit.[35]

In formal terms this idea can be expressed in terms of a refined principle of economy applied to trees. We assume, following Hoekstra and Gueron (1988), that the 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 c-command as the default form in which the morphological affix -ed (pronounced as <u>did</u>) <u>directly</u> c-commands the verb (in an x-chain).   Lasnik (pc) has argued that <u>did </u>is the spellout of a past tense Feature. Therefore we have in effect <u>Feature-command</u>:

34) C-command should be morphologically direct.

This can be illustrated in tree-form. In (29a) the T (tense node) dominates a V which dominates another T, while in (29b) T dominate T directly.

35)a.   TP

---

[35]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.

```
              /       \
         Spec        T_X
         /         /    \
        /         V      NegP
     / /   \      |    \
    /      V   T_X    |     \
   /       |    \     |       \
  you     talk   ed   Neg      VP_X
                       |        |
                      not      t_X
```

35) b. TP
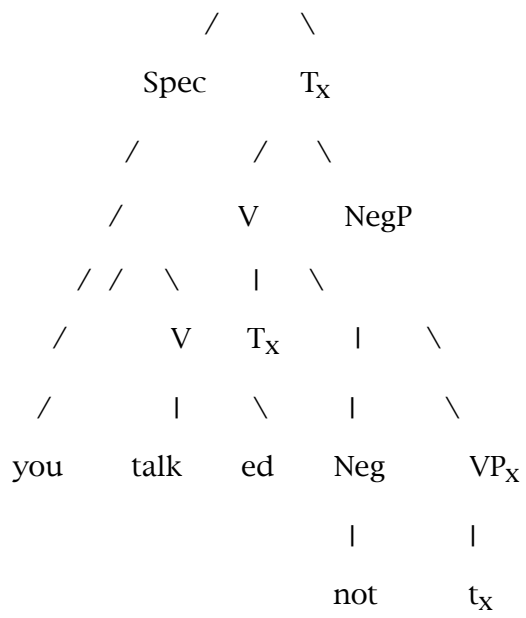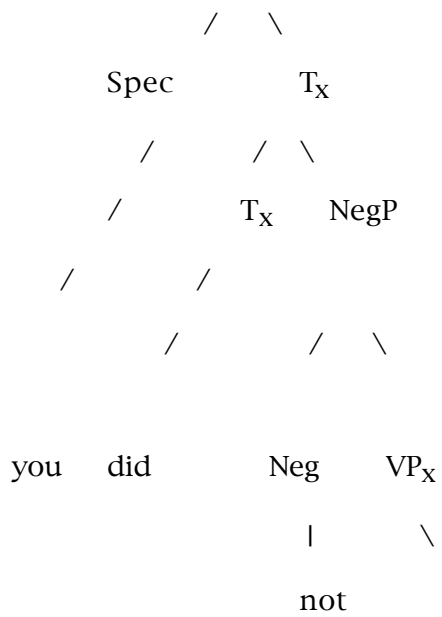
```
            /    \
    Spec          T_X
      /         /  \
     /        T_X    NegP
   /         /
         /          /   \



you    did        Neg     VP_X
                   |        \
                  not


    <==covert============ talk
```

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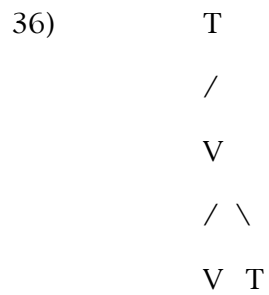
36)   T
              /
              V
             / \
            V   T

As opposed to a direct link (37):

```
37)    T

       |

       T
```

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.[36] 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)  T                          T

       /                          /

      V         =>              T

    /  \                      /  \

    V   T            V      T

    talk  ed                talk    ed
```

---

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

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 <u>wonder</u>.

There are, in fact, a variety of technical options for refining the Feature-checking mechanism. Our goal here is simply to argue that <u>did talk</u> is simpler than <u>talked</u> 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 <u>do</u>-insertion forms. Exactly this occurs in both English and Swiss German (see also Penner (1994). Thus we find (without any emphatic stress) (39) and tense-doubled forms (40):

    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"

40) b. "I did broke it"[37]

"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.

4.1 "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)

---

[37]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.

"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.

Conversely, only forms defined within MDG will overgeneralize beyond lexical classes. [38]   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:

---

[38]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).

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 Hirschensohn (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.

4.2 Incompaltible 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 (1997) summarizes 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 seem 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, PP's, and topicalized NP's.   (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:

       46). *"nothing" did John say

Why is auxiliary inversion insufficient? In contrast, question formation and locative inversion with polarity items obligatorily requires inversion, but <u>only</u> 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 (1997).[39]

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.

---

[39]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 <u>be</u> raises in English, then is it evidence for V2 or residual V2? The answer depends on whether <u>be</u> itself is seen as a Main Verb or an auxiliary.

We will sketch an analysis of each form.[40] 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 <u>do</u> 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 (<u>no one</u>) or wh-word. Therefore inversion is obligatory in cases like:

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.[41] Temporal anchoring is a form of Specificity of the same kind that is indicated for NP's or DP's.

---

[40]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.
[41]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 adjudged 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.

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[42]

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
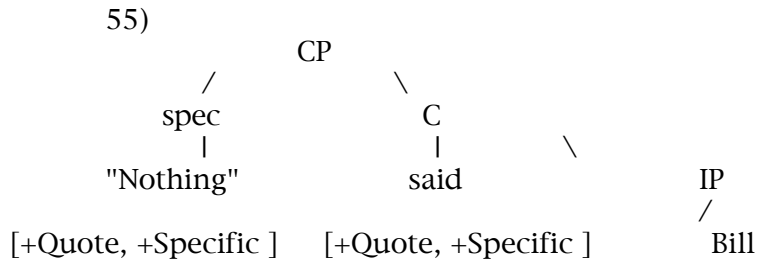
(Bill pushed Bill's car)

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 ?[43] Thus we have:

---

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

[43] The temporal anchoring property provides an explanation to what Collins says is a stipulation in his theory:

55)
```
                        CP
                  /            \
              spec              C
               |               |        \
           "Nothing"          said        IP
                                        /
                                       Bill
      [+Quote, +Specific ]    [+Quote, +Specific ]
```

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 it requires

the elimination of a more economical representation

in either grammar.

This is like the suggestion by Yang (Yang (1999)) that local maxima exist which are incompatible, but since each receives sufficient support, they remain in a "steady state." [44]

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

---

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

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.[45]

5.0 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, 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:

---

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

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 <u>do</u>-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?

5.1 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.[46]   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.

---

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

Suppose further:

61) No LF operations occur inside words,

Therefore morphological economy requires an <u>adjacent, linear array</u> 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: <u>destructiveness</u> versus *<u>destructnessive</u>. Baker (1988) has argued that similar constraints hold for syntactic morphemes.[47] 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 <u>do</u>-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

---

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

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.[48]

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.

 5.2 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 of 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.

---

[48]These are the formal options, but the reality is more complex. The presence
of *passe compose*   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.

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 PP's, then the expression whereafter constitutes, in miniature, a different grammar.[49] We leave this speculation as a suggestion which should be addressed in terms of a richer theory of speech register variation.

6.0 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:

---

[49]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 (1990)) 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. deVilliers (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[50]

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

---

[50]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.

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.

6.1 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.

7.0 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.

BIBLIOGRAPHY

Abdul-Karim, L. (1996) "The Acquisition of Case and Ellipsis"

Paper presented at IASCL, Istanbul, UMass. ms.

Baker (1988) Incorporation: a theory of grammatical function-changing
        University of Chicago

Berwick, R. (1985) *The Acquisition of Syntactic Knowledge*, MIT Press

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

Bickerton, D. (1975) Dynamics of a Creole System Cambridge Cambridge
        University Press

Budwig, N. (1989)   "The linguistic marking of Agentivity and Control
        in Child Language" JCL 16, 263-284

Cavar, D. and C. Wilder (1996) "Auxiliaries in Serbo-Croatian and English" ms.
Potsdam

Collins, C. (1991) "Why and How Come"
        MIT ms.

Chomsky, N. (1986) Knowledge of Language Praeger

Chomsky, N. (1989. 1991) "Some Notes on Economy of Derivation and Representation"
        in R. Frieden ed. Principles and Parameters in Comparative Grammar
Cambridge   MIT Press

Chomsky, N. (1995) The Minimalist Program MIT PRESS

Chomsky, and M. Halle (1968) The Sound Pattern of English

      Harper and Row

Davis, Henry 1987. The Acquisition of the English Auxiliary System and its Relation

to   Lingiustic Theory   Dissertation   University of British Columbia

deVilliers, J. (1991) "Why Questions" in T. Maxfield and B. Plunkett *op, cit.*

Jang, Youngjun (1995) "Against LF Category Movement"

      ms. Harvard University

Haegeman, L. (1990) "Understood subjects inEnglish diaries"

      Multi-lingua, 157-199

Hakannson, G. (1998) "Language Impairment and the Realization of Finiteness"

      eds. A. Greenhill, M. Hughes, H. Littlefield, and H. Walsh,

      Proceedings of the 22nd Annual BU Conference on Language Development

      Cascadilla Press

Hirschensohn, J. (1998) "Minimally Raising the Verb Issue"

      Proceedings of the 22nd Annual BU Conference on Language Development

      eds. A. Greenhill, M. Hughes, H. Littlefield, and H. Walsh, Cascadilla Press

Hollebrandse, B. and T. Roeper (1996) "The Concept of Do-Insertion and the
Theory   of INFL in Acquisition" <u>Proceedings of GALA</u>
eds. C. Koster and F. Wijnen, Centre for Language and Cognition, Groningen

Hoekstra, T. and J. Gueron (1994) "The temporal interpretation of rediction" In ed. A.
Cardinaletti   <u>Syntax and Semantics 28</u> Academic Press

van Kampen, J. (1996) "PF/LF Conversion in Acquisition" in ed. K. Kusumoto
<u>Proceedings of NELS 26</u>, UMass GLSA

Kayne, R. (1994) <u>The Antisymmetry of Syntax</u> MIT Press

Kroch,A. and A. Taylor (1997) "Verb Movement in Old and Middle English: dialect
variation and Language Contact" in van Kemenade, A. and N. Vincent (eds.)
<u>Parameters of Morphosyntactic Change</u> Cambridge: Cambridge University
Press
Labelle, M. (1990) "Predication, Wh-Movement and the Development of Relative
Clauses" <u>Language Acquisition 1</u>, 95-118

Lebeaux, D.(1988) *Language acquisition and the form of the grammar.* PhD
Dissertation, U.Mass Amherst.

Lebeaux, D. 1990. "The Grammatical Nature of the Acquisition   Process: Adjoin-a
and   the Formation of Relative Clauses" in   L. Frazier and J. deVilliers *Language
Processing and Language Acquisition*   Kluwer

Loeb, D. and L. Leonard (1991) "Subject Case-Marking and Verb   Morphology in

normally developing and specifically language-impaired children" JSHR 34,

340-346

Maxfield, T. and B. Plunkett (1991) Special Issue on the Acquisition wh-

UMass Occasional Papers in Linguistics, Amherst, Mass. GLSA

Müller, N. (1993) Komplexe Sätze: Der erwerb von COMP und von

Worstellungsmustern bei bilingualen Kindern Gunter Narr

Müller, N. (1993) "Transfer in Bilingual First Language Acquisition"

Bilingualism: Language and Cognition 1 (3), 1998, 151-171

Marcus, G. et al (1995) "Do Overregularizations come from a grammatical

reorganization" UMass, ms.

Meisel, J. (1994) Bilingual First Language Acquisition: French and German

Grammatical Development Amsterdam: Benjamins

Meisel, J. and M.J. Ezeizabarrena (1996) "Subject-verb and Object-verb Agreement

in   early Basque" in Generative Perspectives on Language Acquisition

ed. H. Clahsen

Penner,Z. (1994) "Learning-theoretic Perspectives on Language Disorders in

Childhood" ms. University of Bern

Penner, Z. and K. Wymann (1998) Normal and Impaired Language Acquisition.

Studies   in Lexical, Syntactic, and Phonological Development

Perez-Leroux, A. (1995) "Resumptives in the Acquisition of Relative Clauses"
Language   Acquisition 4,1&2

Pinker, S. (1984) Language Learnability and Language Development
        Harvard University Press

Powers, S. (1996) "Early Subjects and Agreement" ms. Potsdam

Randall, J. (1981) "Acquisition of -ity and -ness" Journal of Psycholinguistic
Research

 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. (1995) Talk presented at GALA 1995

Roeper, T. (1993) "The Least Effort Principle in Child Grammar: choosing a
        marked   parameter" Knowledge and Language 1: from Orwell's Problem to
Plato's   Problem eds.W. Abraham and E. Reuland

Roeper, T. and J. deVilliers (1992) "The One Feature Hypothesis"
        UMass ms.

Roeper, T. 1996 "The Role of Merger Theory and Formal Features in

        Acquisition" in eds. H. Clahsen and R. Hawkins <u>Generative</u> <u>Perspectives on</u>

<u>Language</u>   <u>Acquisition, Empirical Findings, Theoretical Considerations, Cross-</u>

        <u>linguistic Comparisons</u>

Roeper,T. and B.Rohrbacher (1994) "Null subjects in Early Child   English and the

theory of Economy of Projection"

        Paper presented at Bern, Switzerland, January.

Roeper, T. and A. Perez (1997) "The Interpretation of Bare Nouns in   Semantics

        and Syntax: Inherent Possessors, Pied-Piping, and Root   Infinitives" in ed.

J.   Schaeffer, J. (1997)   <u>Bare Nouns and Root Infinitves</u> MITWPL

Rubin, E. (1996) Talk on Bilingualism and Minimalism, UMass

 Schafer, R. and T. Roeper (1999)"The Role of the Expletive in the Acquisition of a
Discourse Anaphor" ms. UMass

Speas, M. (1994) "Null Arguments in a Theory of Economy of Projection" in E.

Benedicto   and J. Runner (eds.)*Functional Projections* UMOP 17

Schütze, K. and K. Wexler (1996) "Subject Case Licensing and English   Root

        Infinitives" BUCLD eds. A. Stringfellow, D. Cahana-Amitay,

        E. Hughes and A. Zukowski <u>Cascadia Press</u>

Stromswold, K. (1995) "The acquisition of subject and object questions" <u>Language</u>

        <u>Acquisition 4.1&2</u>

Takahashi, M.(1989) "Object Inversion in Wh-questions" UMass ms.

Tracy, R. and E. Lattey (1994) How Tolerant is Grammar Niemeyer

Tracy. R. (1996) Child Languages in Contact: Bilingual Language Acquisition (English/German) in Early Childhood Habililtationsschrift Tübingen

Vainikka, A. (1994) "Case in the Development of English Syntax" Language Acquisition Vol. 3,3

Vainikka, A. (1990) "The Status of Grammatical Default Systems" in   L. Frazier and   J. deVilliers Language Processing and Language Acquisition

Wexler, K. (1998) "Optional Infinitives, Tense, and Unique Checking" (to appear) Special Issue of Lingua

Wexler, K. and P. Culicover (1980) Formal Principles of Language Acquisition MIT Press

Yang, C. (1999) "The Variational Dynamics of Natural Language: Acquisition and Change" MIT ms. Artificial Intelligence Laboratory

Zwart, J. (1993) Dutch Syntax: a minimalist approach. PhD   dissertation, Groningen