UCLA Working Papers in Linguistics, Papers in Mathematical Linguistics 1
Volume 15, Article 1: 1–17, 2010

On Language Variation and Linguistic Invariants[1]

Edward L. Keenan and Edward P. Stabler

Keywords linguistic universals, linguistic invariants, language variation

We present a brief, informal overview of a formal approach to grammatical variation developed in K&S (Keenan and Stabler 2003), to which we refer the reader for proofs and other formal statements. Our purpose here is to show how we can formulate structural universals of grammar given grammars that are structurally distinct in relevant respects.

1 The One and the Many

A central quest in syntactic theory is to reconcile the audible diversity of natural languages (NLs) with the claim that they have a common, biologically determined, form.

1.1 The One

Mainstream syntactic theory (MST) from Chomsky (1957) to Bošković and Lasnik (2007) attempts this reconciliation by building it into the form of individual expressions which must satisfy general constraints on rules/derivations and representations. For a *core* expression X of L, MST asks "What is the structure of X?" The initial response, now, is often a binary branching tree in Spec-[Head-Complement] order, using language independent category symbols and structure building functions (Merge, Move). Regarding variation, some, perhaps much, is relegated to the *periphery* beyond *narrow* syntax and ignored; some is acknowledged in parameters with small ranges (e.g. question words remain in situ or front); and some lies in feature variation forcing slightly different patterns of movement/copying. Overt morphology is language specific, not determined by UG and not structurally autonomous (Bobaljik 2002) but a "reflection" of hierarchical constituent structure.[2]

[1]This paper is a slightly augmented version of an LSA plenary session presentation by the first author on January 5, 2007.

[2]Borer (2005a,b) is something of an exception to this claim. And in general imputation of properties to MST are not intended to hold exceptionlessly for those who contribute to that tradition. Note too that features called morphological can be checked without requiring morphology to be overt.

So MST focuses on the unity pole of the unity-diversity continuum, treating different languages as syntactically and semantically similar.

1.2 *The Many*

K&S in contrast, proposes a reconciliation that focuses on diversity. It provides a conceptual notion of structural invariant that can be satisfied by non-isomorphic structures. It formalizes the early Chomkyan desideratum that linguistic operations are *structure dependent* (Chomsky 1965: 55–56; 1975: 30–35; Radford 1997: 11–15). Often invariants are not present in the grammar as conditions on rules or representations, and are not instantiated by single expressions. Given an expression X, we ask “What expressions have the *same* structure as X?” (not “What is *the* structure of X?”). Perhaps some feeling for our relational approach to linguistic invariants can be given with the following, very imperfect, anthropometric analogy: The absolute height of human is certainly not invariant: it varies from about 2 to 8 feet. But the ratio of arm span to height is much closer to being invariant (the same). Also, in presenting our perspective we alert the reader to one difference in character between it and MST:

MST focuses on the form of syntactic theory.

Minimalist principles such as inclusiveness and economy are notation based — they constrain how we derive expressions, not directly what can be derived. Kayne (1994) opens with the central role of notation (emphasis ours): “It is difficult to attain a restrictive theory of syntax. One way. . . is to restrict the space of available syntactic *representations*, for example, by imposing a binary branching requirement,. . . The present monograph proposes further severe limitations on the range of syntactic *representations*. . . ”. But restricting notation can be detrimental. Different notations may suggest different questions and generalize differently to new phenomena. For example, in the early 1960’s context free grammars and categorial grammars were shown to define the same class of languages (Bar-Hillel, Gaifman, and Shamir 1960). But it was categorial grammar that most naturally captured function-argument structure (Lewis 1970; Montague 1973), effectively creating the field of formal semantics as we know it today. The *significant* properties of objects are invariant under changes of descriptively comparable notation, notational artifacts aren’t. The truth of *x is hotter than y* does not vary according as we measure temperature in Fahrenheit or Celsius. The truth of *x is twice as hot as y* does. So let us focus on regularities of linguistic nature, not their notational expression, adopting the slogan

If you can’t say something two ways you can’t say it.

1.3 *Back to the Many*

Here first is a semantic generalization in propositional logic based on notationally distinct objects. (1a) uses standard infix notation for conjunction and disjunction, (1b) uses prefix (Polish) notation. $\models$ is “entails”.

(1) a. $((P \vee Q) \wedge \neg P) \models Q$
b. $\wedge \vee PQ \neg P \models Q$

The notational variation is non-trivial: prefix notation uses no parentheses and has no structural ambiguities. Eliminating all parentheses from infix notation yields structural and semantic ambiguity: $((P \wedge Q) \vee R) \not\equiv (P \wedge (Q \vee R))$. On the other side, students of logic tend to find infix notation easier to understand, especially in long formulas. But the entailment fact in (1a,b) is *the same*: a disjunction of formulas conjoined with the negation of one entails the other. The entailments expressible in the two notations are the same. The synonymy of $(P \wedge Q)$ and $\wedge PQ$ is determined by independent compositional interpretation. We have no need to derive them from a common "deep structure" or map them to a common "LF".

2 On Structure

We treat a language as a compositionally interpreted set of expressions defined by a grammar $\mathrm{G} = (Lex_{\mathrm{G}}, Rule_{\mathrm{G}})$, where Lex_{G}, the *lexicon* of G, is a finite set of expressions called *lexical items* and $Rule_{\mathrm{G}}$ is a set of *structure building functions* called *rules*, which iteratively map tuples of expressions to expressions. The *language* L_{G} *generated* by G is the set of expressions derived from Lex_{G} by finite iteration of the rules. So L_{G} is the closure of Lex_{G} under $Rule_{\mathrm{G}}$. A *descriptively adequate* G for a NL L is *sound* (everything it generates is judged by competent speakers to be in L) and *complete* (everything speakers accept is generated).

A grammar G is *isomorphic* to a grammar $\mathrm{G}', G \cong G'$, iff there is a bijection $h : L_{\mathrm{G}} \rightarrow L_{\mathrm{G}'}$ matching the functions F in $Rule_{\mathrm{G}}$ with the F' in $Rule_{G'}$ so that when F derives an expression z from some expressions $x_1, \ldots, x_n$ then F' derives $h(z)$ from $h(x_1), \ldots, h(x_n)$, and conversely. An isomorphism h from G to itself is called an *automorphism* (or *symmetry*) of G. They are the ways of substituting one expression for another within a language which do not change how expressions are built. So they preserve structure: F builds z from $(x_1, \ldots, x_n)$ iff F builds $h(z)$ from $(h(x_1), \ldots, h(x_n))$, all automorphisms h. The set $\mathrm{Aut}_{\mathrm{G}}$ of automorphisms of G represents the "structure" of G. It contains the identity map, and it is closed under function composition and inverses, so is a group, the *automorphism* (or *symmetry*) *group* of G, noted $\mathrm{Aut}_{\mathrm{G}}$ or $\mathrm{Sym}_{\mathrm{G}}$.[3]

2.1 *Complexity*

In cases of interest (natural languages) L_{G} is infinite. Since automorphisms map an infinite set to itself one might think that there are massively many of them, and thus hard to study and characterize in any given language. But in fact in cases of interest the number of automorphisms of G is finite. The reason is that the value an automorphism h assigns to a derived expression $F(x_1, \ldots, x_n)$ is uniquely determined by the values it assigns to the expressions it is built from, namely $x_1, \ldots, x_n$. So

[3] Rules are partial functions, so we define $g \circ f$ as that map with domain $\{a \in \mathrm{Dom}\, f \mid f(a) \in \mathrm{Dom}\, g\}$ given by: $(g \circ f)(a) = g(f(a))$.

once we have given the values of an automorphism on the lexical items— finite in number in cases of interest— we have determined its values at all expressions: $h(F(x_1,\ldots,x_n)) = F(h(x_1),\ldots,h(x_n))$. ($\mathrm{Aut}_G$ is *forced* to be finite if Lex_G is finite and none of its elements is also a derived expression).

2.2 *Invariants*

An expression s has the same structure as an expression t iff $h(s) = t$ for some automorphism h. We do not mention here "the structure" of an expression— an epistemological plus, as different syntactic theories agree more readily that *John sang* and *Bill danced* have the same structure, than they do about what the structure of *John sang* is. We define:

Definition 1. A relation R on L_G is **invariant** iff $h(R) = R$, all automorphisms h. That is, replacing all tuples $(s_1,\ldots,s_n) \in R$ by $(h(s_1),\ldots,h(s_n))$ leaves R unchanged. So $w \in L_G$ is **invariant** iff $h(w) = w$, all automorphisms h.

Thesis. $w \in L_G$ is a *grammatical formative* ("function word") iff $w \in Lex_G$ and w is invariant.

So function words are items that are isomorphic only to themselves. Replacing them with something else changes structure (usually destroying it yielding ungrammaticality). In the models of grammars in K&S, reflexive pronouns, case, voice, and agreement markers are provably invariant. We expect that lexical invariants correspond to heads of functional projections in more usual terminology. What is new here is a characterization of what counts as *functional*— namely, being a linguistic object (including bound morphemes) which can only be mapped to itself by the automorphisms, the structure preserving functions, of the grammar).

As an item that a function f maps to itself is called a *fixed point* of f, we characterize the functional expressions in a grammar as the fixed points of the syntactic automorphisms. We turn now to some generalizations built on this notion of invariant.

2.3 *Degrees of Invariance, Language Change*

Invariance generalizes to a scalar by:

Definition 2. a. $Inv(x)$, the *degree of invariance* of a linguistic object (expression, property, relation) x is the proportion of automorphisms that fix x, that is, that map x it itself. So if x is invariant per Definition 1, then $Inv(x) = 1$, as all automorphisms map x to itself.

b. x is *more invariant than* x' iff $Inv(x) > Inv(x')$.

These definitions provide a rigorous way to say that conjunctions and prepositions are more grammaticized than intransitive verbs. And we can use this notion to represent the grammaticization (Hopper and Traugott 1993) of an expression w by saying that $Inv(w)$ increases over time. If $Inv(x)$ reaches 1 then x is fully grammaticized.

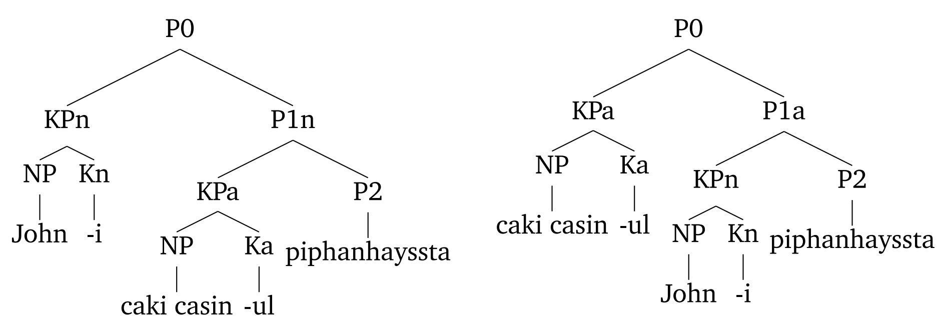
2.4 *Relation Invariants*

K&S prove that for all G (not just G for NLs), *is a constituent of* and *c-commands* are invariant relations. These notions require a more general than usual definition since we have not limited ourselves to rules whose action is modeled by standard labeled trees. Such functions basically just derive expressions by concatenation, not, for example, substitution (widely used to derive consequences from premisses in logical deductive systems for example). Here are some examples of the generalized definitions:

Definition 3. u is an *immediate constituent* of v iff v is the value of a rule F at a tuple $(s_1,\ldots,s_n)$ and u is one of the s_i. u is a *constituent* of v iff u is v or u is an immediate constituent of a constituent of v. s is a sister of t in v iff $s \neq t$ and for some constituent u of v, s and t are immediate constituents of u. s *c-commands* t in u iff s is a constituent of a sister of t in u.

All these relations are provably invariant in all G. In contrast, the property of being an *anaphor* and the relation *x is a possible antecedent of an anaphor y in z* are not invariant in all G, but they are for all G we have constructed to model NLs. This illustrates invariants among non-isomorphic structures. K&S's grammar for minimal clause structure in Korean generates (2b) where anaphors like *self-acc* asymmetrically c-command their antecedents. To see that *self-acc* is an anaphor one must check its semantic interpretation in K&S (see below): it maps a binary relation R to the property $\{a|aRa\}$).

(2)



a. John-nom self-acc criticized
'John criticized himself'

b. self-acc John-nom criticized
'John criticized himself'

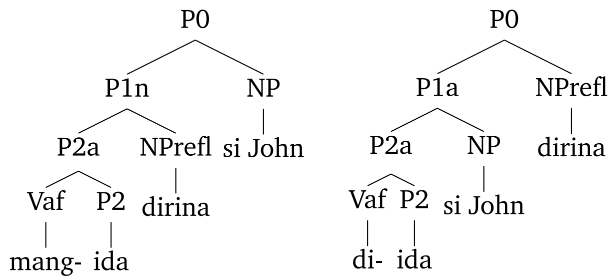
Each of these Ss is interpreted compositionally, and they receive the same interpretation: True iff $(j, j) \in$ criticize, j the denotation of *John*. This fact is not different in kind from the fact that $(P \wedge Q)$ is logically equivalent to the structurally distinct $\neg(\neg P \vee \neg Q)$ in propositional logic.

Lovers of trees Beware! (2a,b) are isomorphic qua ordered labeled trees (same branching structure, each has distinct labels just where the other does). But the expressions are not isomorphic in our grammar: if an automorphism h mapped *john* to *self* and *-nom* to *-acc* it could map *John -nom laughed* to *self -acc laughed,* which is not in our model language, contradicting that h is an automorphism. See K&S p.50.

(3) is from Toba Batak (W. Austronesian; N. Sumatra). See (Schachter 1984; Cole and Hermon 2008). The distribution of anaphors in Toba, as in other W. Austronesian

languages (Tagalog, Malagasy, Balinese) is conditioned by verb voice, the reflexive often occurring as what we thought was a “subject”. And as in the Korean case, lexical items occurring in both Ss have identical interpretations. So again anaphors may asymmetrically c-command their antecedents.

(3)



a. mang- see self John
'John sees himself'

b. di- see John self
'John sees himself'

Theorem 4. *The case markers* -nom *and* -acc *and voice markers* mang- *and* di- *are invariant in their grammars in K&S (pp.49 and 70, respectively).*

Thus is morphology “structural” in the *same* sense as constituent structure: both are *preserved by all the automorphisms*.

Theorem 5. *a. The property of being an anaphor is invariant in K&S's model grammars for English, Korean and Toba Batak (pp.36, 50 and 70, respectively)*

b. The Anaphor-Antecedent relation is invariant in K&S's models (pp.53 and 71, respectively), as it is in their model of English, not illustrated here, in which the usual c-command conditions hold.

Theorem 2 builds on two properties which distinguish our approach to anaphora from more standard ones:

P1. *anaphor* is semantically defined, so we are not free to stipulate that himself in English, *caki casin* in Korean, and *dirina* in Toba Batak are anaphors. Rather, the anaphoric expressions in a language are those whose semantic interpretation satisfies our definition (the same for all languages). Often the anaphors in a language cannot be listed as there are infinitely many (all but finitely many being syntactically complex).

P2. The structural means used to identify anaphors and their antecedents are not structurally the same across languages. Nonetheless the anaphor-antecedent relation is, we claim, invariant in each NL grammar (and so universally invariant).

We explain P1 and P2 and illustrate with examples, then formulate P2 as a universal claim U1.

Referentially Autonomous Expressions (RAEs) such as *Zelda*, *most men who Zelda dates*, etc. combine with $n+1$-ary predicates (P_{n+1}s), to form n-ary ones (P_ns). For simplicity we limit ourselves here to P_ns, $0 \leq n \leq 2$. Semantically P_ns denote

n-ary relations over whatever domain E is under consideration.[4] RAEs denote *Type-1* functions — they map $n+1$-ary relations to n-ary ones, thus reducing arity by 1. But not just any Type-1 function is a possible RAE denotation. The value such a function assigns to a binary ($n+1$-ary) relation is determined by the values it assigns to the unary relations (subsets of E). Thus *Dana praised every student* is true iff Dana is in the denotation of *praised every student*, the set of objects that praised every student, that is, the set of objects b such that *(every student)* holds of the set of things that b praised. For R a binary relation write bR for $\{d \mid (b,d) \in R\}$. So bR is the set of objects d that b stands in the relation R to. Then the possible RAE denotations F are those Type-1 functions satisfying (4):

(4) For R a binary relation, $F(R) = \{b \mid F(bR) = 1\}$.[5]

Theorem 6. *The Type-1 F satisfying (4) are just those which satisfy the AEC (Keenan 1988):*

> **Accusative Extensions Condition (AEC)** *For all $x, y \in E$, all binary relations R, S, if $xR = yS$ then $x \in F(R)$ iff $y \in F(S)$.*

So *most men who Zelda dates* satisfies the AEC since, for example, whenever Sue distrusts just the people that Ann likes then *Sue distrusts most men Zelda dates* and *Ann likes most men Zelda dates* must have the same truth value (both true, or both false).

Now observe that expressions like *himself, herself*, etc. as they occur in *No model hates herself* fail the AEC. If Sue distrusts just Jean, Robyn, Amy, Ann, Mary, and Pat and those are exactly the people that Ann likes then *Sue distrusts herself* is false and *Ann likes herself* is true. Such expressions do however satisfy a weaker invariance condition (Keenan 1988):

Accusative Anaphor Condition (AAC) For all $x \in E$, all binary relations R, S, if $xR = xS$ then $x \in F(R)$ iff $x \in F(S)$.

So if x bears R to the same things that x bears S to then x bears R to himself/herself iff x bears S to himself/herself. And we may, on first pass, define anaphoric expression by:

Definition 7. An expression α in L_G is an *anaphor* iff the interpretation of α satisfies the AAC in all models and fails the AEC in some.[6]

One checks by example that the underlined complex expressions in (5) are anaphors:

(5) a. John criticized <u>every student but himself</u> / <u>no student but himself</u>

---

[4] Note that P_0s denote subsets of $E^0 = \{\emptyset\}$, so a P_0 denotes an element of $\{\emptyset, \{\emptyset\}\} = \{0, 1\}$, our usual representation of the set of two truth values.

[5] Interpreting b as an n-tuple and bR as $\{d \mid (b,d) \in R\}$, the equation in (4) is the general condition for F a map from P_{n+1}s to P_ns.

[6] This definition must be generalized to account for anaphors in a wider range of contexts: *Mary protected John from <u>himself</u>, every worker's criticism of <u>himself</u>*, etc.

b. John criticized only himself and the teacher /neither himself nor the teacher

c. John knows many people smarter than himself /no one as verbose as himself

d. He nominated someone other than himself /He wouldn't nominate anyone other than himself

3 Structural Universals

We consider some candidates U1 — U5 as structural universals of human language. Our purpose is to illustrate how such claims can be formulated in our framework. We think they are plausible, but much empirical investigation is needed.

U1. The property of being an anaphor is a structural invariant of human language.

U1 says that anaphors are always mapped to anaphors by the syntactic automorphisms. Note that even if each lexical anaphor is mapped to itself, each complex anaphor typically will not be. An automorphism of English might map *himself* to *himself* but map *no doctor but himself* to *no lawyer but himself*. Nonetheless, U1 says that in each L_G anaphors have some syntactically distinctive properties, ones which may fail to be comparable across languages. In nuclear clauses in Batak for example they concern distributional constraints with respect to *mang-* vs. *di-* prefixed verbal roots. In Korean they concern case marking, and in English it is Principle A that distinguishes the distribution of anaphors: *John laughed* is fine, *Himself laughed* is not. These same observations support U2:

U2. The relation *x is a possible antecedent of an anaphor y in z* is a linguistic invariant.

3.1 Theory Design

Merely representing a NL as a pair $(Lex_G, Rule_G)$ is not a theory — at best it is a common denominator of theories such as Minimalism, HPSG, LFG, RG,... But two features of our approach do have a liberating effect on theory design: (i) the requirement of a compositional semantics, and (ii) the theorem that morphology may be structural. From (ii), in designing a grammar we are free to condition the distribution of anaphors (semantically defined) directly in terms of case or voice, it is not necessary to derive them from, or reconstruct them to, forms c-commanded by their antecedents.

From (i), semantic representations — on our view compositionally interpreted audible structures — differ from language to language, whereas in MST (Higginbotham 1985, Chomsky 1986:156, Hornstein 1995:§1) LFs for different languages are assumed or argued to be roughly the same. They claim that primary semantic data do not suffice for the child to infer the structure of LF parameters (in distinction to primary phonetic data which do permit the inference of diverse phonologies).

Since we do learn to use language meaningfully the semantic module must be innate, hence essentially the same across speakers of different languages.

The argument does not convince.

No characterization of primary semantic data is given, nor are reasons why they are insufficient to set "LF parameters". In fact children learn to use language meaningfully as they learn to pronounce it. Simple situations in which they follow commands, make requests, answer questions, disagree,... are easily seen to contribute to learning the meaning of expressions — reference (count and mass), basic argument structure and theta roles, modification (*Hand me the red crayon.* — *No, no, not the* blue *one, the* red *one!*), etc. And of course learners of different languages are interpreting different expressions, an unproblematic fact as our example with $((P \vee Q) \wedge \neg P)$ and $\wedge \vee PQ\neg P$ shows. (6) and (7) below exhibit pairs of logically equivalent sentences with different internal structures — showing that different structures may determine the same (logical) meaning.

(6) a. Between a third and two-thirds of Americans brush their teeth regularly
 b. Between a third and two-thirds of Americans don't brush their teeth regularly

(7) a. All but two students read at least as many poems as plays
 b. Exactly two students read more plays than poems

(6) is surprising as they differ just in that the predicate in (6b) negates that in (6a). The pattern is general as long as the fractions sum to 1 (Keenan 2004). We should note that LF is not designed to represent meaning in general. It doesn't even support a definition of entailment (Chomsky 1986:67n11) and several of the semantic notions it does represent are esoteric and likely learned later than the notions we mentioned above. For example, relative scope of quantifiers only arises with two quantified arguments of a given predicate. Languages provide many means for distinguishing the arguments of transitive verbs (word order, agreement, case marking, chain of being orders) but do not systematically disambiguate scope (Keenan 1988). T. Lee (1986) supports experimentally that both English and Chinese children understand some basic quantifiers in intransitive Ss by age 4, but even by age 8 do not have adult competence on their relative scope judgments in transitive Ss. Even in formal logic it was only with Henkin (1961) that we learned to construct formulas with universal and existential quantifiers lacking scope dependencies.

3.2 Property Invariants

Is the property of being a lexical item invariant in all G for NL? That is, do automorphisms always map lexical items to lexical items? This is a possible empirical truth (not a theorem) even though the lexicons of different NLs differ. Similarly, is the property of having a given grammatical category invariant? That is, do all automorphisms map each phrase of category C to a phrase of category C? Taking this as axiomatic would provide a universal structural role for categories. But K&S show this fails: for example, it is possible to design a grammar in which an automorphism exchanges the masculine and feminine nouns and adjectives.

Another idea is that a category C is an equivalence class of expressions defined by a coarsest congruence, where a *congruence* is an equivalence with the property that for every rule F that applies to expressions $(s_1,\ldots,s_n)$, if $(t_1,\ldots,t_n)$ is such that s_i is equivalent to t_i, all $1 \leq i \leq n$, then F applies to $(t_1,\ldots,t_n)$ and $F(s_1,\ldots,s_n)$ is equivalent to $F(t_1,\ldots,t_n)$. But this fails too for some reasonable grammars. It rules out, for example, grammars with certain kinds of identity conditions. For example, a rule that allowed coordination of any two *distinct* NPs— *both John and Bill* but not *both John and John*—would treat *John* and *Bill* as the same category even though one cannot generally replace the other in the domain of coordination.

K&S argue instead for the weaker U3. U3 provides a universal structural role for categories which allows grammars of different languages to have different categories. Embarrassingly the field has no purely syntactic definition of category that allows us to infer that category C in language L and category C' in language L' are the same. See Baker (2003) for some discussion.

U3. For all stable automorphisms h, all expressions x, if x has category C so does $h(x)$.

Definition 8. a. An automorphism of a grammar G is *stable* iff it extends to an automorphism of each lexical extension of G.

b. G_n is a *lexical extension* of G iff there is a sequence $(G_1,\ldots,G_n)$ of grammars such that each G_{i+1} differs from G_i just by the addition of a single lexical item isomorphic to one in Lex_{G_i}. (The set of stable automorphisms in Aut_G is always a subgroup of Aut_G).

K&S's model of Spanish for example has two gender classes of nouns, Nm and Nf, and overt agreement of adjectives and dets with Nouns. There are automorphisms h which interchange the lexical Nms and Nfs, provided these two sets have the same number of members. If we add one new member to just one of the classes we can no longer interchange them by an automorphism, so such h are *unstable*: their status as automorphisms is not preserved under trivial additions to the lexicon. A fourth candidate for a structural universal is:

U4. Theta role assignment is invariant.

So if x bears theta role τ to y in z then $h(x)$ bears τ to $h(y)$ in $h(z)$, h any automorphism. That is, theta role assignment Θ is a function of structure. If *Ed* has different theta roles in *Ed ran* and *Ed arrived* then these Ss must be non-isomorphic (per current theories). But *Ed* may occur in structurally distinct environments and still be assigned the same theta role. U4 is strictly weaker than UTAH, which requires that Θ be (structurally) one to one: so same theta role $\Rightarrow$ isomorphic sources. But functions may take the same values at different arguments: $(2+3)=(1+4)$. So active subjects and their corresponding passive agent phrases may originate in non-isomorphic configurations.

3.3 Greenberg Duality

We close with a much more contentious candidate universal, U5 below. Two languages are word order *duals* if the expressions of one are the mirror images of those of the other. Lexical items are self dual. A rigid SXOV L, like Turkish, is (isomorphic to) the dual of a rigid VOXS language, like Malagasy. Formally, the dual v^d of a sequence $v = (v_1, v_2, \ldots, v_n)$ of lexical items is just its mirror image, $(v_n, \ldots, v_2, v_1)$. The dual K^d of a set K of expressions is the set of w^d for $w \in K$. If F is in $Rule_{\mathrm{G}}$ then its dual F^d is that function with domain $\mathrm{Dom}(F)^d$, that maps $(w_1^d, \ldots, w_n^d)$ to the dual of $F(w_1, \ldots, w_n)$. We define G^d to be that grammar with the same lexical items as G and whose rule set is the set of duals of rules of G. We have:

Theorem 9. *a.* $(L_{\mathrm{G}})^d = L_{\mathrm{G}^d}$

b. $\mathrm{G} \cong \mathrm{G}^d$*, the map sending each* $w \in L(G)$ *to* w^d *is an isomorphism.*

Theorem 9a just says that the dual of the language is the language generated by the dual grammar, so we defined G^d right. Theorem 9b says that G and G^d are isomorphic.

U5. The set PHG of possible human grammars is closed under isomorphism.

U5 just says that if $G \in \mathrm{PH}_{\mathrm{G}}$ and $\mathrm{G} \cong \mathrm{G}'$ then $G' \in \mathrm{PH}_{\mathrm{G}}$. Our justification for U5 is that UG only selects for structure not content and thus cannot distinguish between isomorphic variants (though other constraints, say ones on possible phonological systems, might rule out some isomorphic images as being non-natural on other grounds).

Corollary (Duality): PH_{G} is closed under duals. (From U5 and Theorem 9b).

The Duality Corollary makes us hesitant to accept Kayne's Antisymmetry axiom, which forces right branching structures. If only such grammars were acceptable then PH_{G} would not be closed under duals. But since there are left branching Ls (Toba Batak, Malagasy) the Antisymmetry axiom must be weakened.

But U5 is also problematic. The Duality Corollary suggests an equal distribution of word order types and their duals, which is not the case. Right branching (SXOV) Ls are the most common across areal and genetic groupings whereas VOXS Ls are a clear minority, but include Malagasy (Keenan 1976) and several other Austronesian languages, and Tzotzil (Aissen 1987) and several other Mayan languages. Worse, the OVS duals of SVO languages have (to our knowledge) just Hixkaryana (Carib; Brazil) (Derbyshire 1977) as a well attested exemplar, while the OSV duals of VSO Ls are just barely attested: *Ethnologue* (Gordon and Grimes 2005) cites Jamamadi, an Arawakan language in Brazil.

But immediate rejection of U5 would be short sighted. Much work in the physical sciences, esthetics, mathematics, and the philosophy of science supports both the fundamental role of symmetry in the phenomena under study and also the presence of asymmetries and spontaneous symmetry breaking. Acknowledging and studying these asymmetries and symmetry failures has been significant stimulus to deeper understanding. "The study of anomalies now plays an important role in our search for the symmetries of nature" (Zee 1986: 300).

We can hardly summarize here the basic, and at times dazzling and provocative, work in this area. We just note a few highlights that have influenced our thinking and point the reader to several accessible and enlightening introductions to symmetry and symmetry breaking. Weyl (1952) is a classic, discussing symmetries, and asymmetries, in physics, biology, and art. Bunch (1989) and Gardner (2005) are more recent and very informative. Darvas (2007) focuses more on symmetry in art, and Zee (1986) on symmetry in physics. On the mathematical side we note Stewart and Golubitsky (1999) plus any basic textbook on group theory, the language of symmetry and invariance, for example Rotman (1999: §1–3). On the more philosophical and epistemological side we have found van Fraassen (1989) and Wigner (1979) enlightening.

Concerning the conceptually fundamental nature of symmetry and invariants (what remains unchanged under the action of the symmetries) the first author must acknowledge his awe at Felix Klein's Erlangen dissertation (Klein 1893) in which he stood Euclidean geometry on its head. The objects of study became whatever was invariant under the action of translations, rotations, and reflections. Other geometries are invariants of other transformations. Later topology, which grew out of geometry, became the study of continuous transformations and topological invariants those objects, properties, . . . which remain invariant under these transformations. In 1918 Emmy Noether (Bunch 1989:95; Brewer and Smith 1981; Cole 1997:183), at Erlangen, proves that symmetry principles in physics (including relativity theory) imply conservation laws. Indeed Hermann Weyl (cited in Bunch 1989:144) claims "The entire theory of relativity. . . is but another aspect of symmetry". Gardner 2005: 337 quotes Einstein to the effect that invariant theory would have been a better name for his achievement than relativity. More recently symmetries have been used to predict new elementary particles (Sternberg 1994). And symmetry breaking is the exclusive subject of a recent monograph (Strocchi 2005). Still, "Why", asks Feynman, "is nature so nearly symmetrical?" (Bunch 1989: 189).

Cotton (1990), a standard textbook, exemplifies the utility of symmetry in chemistry by using group theory to classify molecules by structure: ". . . the number and kinds of energy levels that an atom or molecule may have are rigorously and precisely determined by the symmetry of the molecule or of the environment of that atom." Cotton 1990: 3. In mathematical domains group theory, as noted, the mathematics of symmetry and invariants, has become a major subfield in mathematics from its initial impetus by Galois (1811–1832). In logic, Tarski (1986) presents informally the idea that "logical operations" are simply the most general ones (in distinction to translations, etc. which must obey constraints in addition to being permutations of the domain. Keenan (2001) studies this idea more formally). And of interest Roman Jakobson (1963) pushed the notion of linguistic invariant in the famous 1963 Universals conference: "Naive attempts to deal with variations without attacking the problem of invariants are condemned to failure" (Jakobson 1963: 272).

We return now to the unequal distribution of word order duals — which, being isomorphic but distinct, we might have expected to be equally distributed (perhaps assuming a Leibnizian "Sufficient Reason" basis for Nature). But as we have seen they are not. Nor is right and left spiraling DNA. As far as we know, all animals are built from right-handed DNA, though it seems that a little left-handed DNA has been found in nature, and it can be synthesized. So what accounts for the statistical disparity? For that matter, why do righthanders outnumber lefthanders? In some cases at least we feel the choice was arbitrary, but once established it self-perpetuated. Driving on the left or the right is an arbitrary convention, entailing building cars with the steering wheel on the left or the right. But as one came to dominate in continental Europe Sweden was pushed to "walk in step", that is, drive on the right. Britain is still holding out. (Japan, Australia, and New Zealand continue to drive on the left).

Now, to return to properly linguistic considerations, how good is the comparison between linguistic symmetries and invariants and those in physical or mathematical fields? Here the comparison is very good indeed. In both cases we study what is preserved under the action of classes of structural functions. So we are not simply making some analogy here, we are using linguistic science as another case where symmetries and their invariants can be studied. But will this approach lead to enlightening results, as it has in the fields discussed above? And here of course we don't know how enlightening this approach will be until we pursue it detail to see where it leads. There are however some grounds for a preliminary positive assessment.

In the first place, we have claimed a methodology that permits the description of structural regularities across structurally non-isomorphic languages. This is already a strong reason to pursue our approach. And secondly, we can benefit from the massive amount of work that has gone into the study of symmetries and the mathematical apparatus (group theory) needed to describe them. Minimally we can ask if there is something distinctive about the automorphism groups (symmetry groups) for grammars of human languages. Can we distinguish human languages from other formal systems by the structure of their symmetry groups? To pursue an answer to such a big question we need to know massively more about the group structure of well studied grammars.

Do any of the following traditionally structural properties of grammars force some distinctive property on the automorphisms of G: paradigm, inflectional morphology, allomorph, subcategory, extraction and copy rules (Kobele 2006), category changing operations, case and voice marking? Are locality constraints or cyclic domains (or phases) characterizable in terms of automorphisms? We don't know. Some simple classes of objects have characteristic automorphism groups. For example regular polygons (n-gons) have dihedral groups (n rotations, n reflections). Do the $\mathrm{Aut}_{\mathrm{G}}$ for natural languages have any characteristic structure? We simply don't know.

But merely modeling the simplest type of agreement phenomena we learned that some automorphisms are unstable, and may fail to extend to automorphisms of grammars that trivially augment the original just by adding a new lexical item isomorphic to an old one. So unstable auts may indicate basic linguistic symmetries that can be overridden by default forms, as in coordination. How much other linguistic information will be coded in the automorphisms? Can we for example reconstruct the grammatical category distinctions among lexical items given the automorphisms? For example we hypothesize that for a given lexical item d, the other lexical items of the same category as d are just those in Orbit(d) under the stable automorphisms, where Orbit(d) is the set of expressions that a stable automorphism can map d to.

Finally, the Corollary invites a deeper, less speculative comparison with MST: eliminating redundancy can be overridden. Among the virtues of axiom systems mathematicians include independence (non-redundancy) — no axiom is to follow from the others. But:

Symmetry trumps redundancy.

For example, common axiomatizations for Boolean algebra contain the two distributivity laws:

(8) a. $(x \wedge (y \vee z)) = (x \wedge y) \vee (x \wedge z)$
b. $(x \vee (y \wedge z)) = (x \vee y) \wedge (x \vee z)$

In (8a) meet $\wedge$ distributes over join $\vee$ and in (8b) join distributes over meet. If (8a,b) are both removed then neither is entailed by the remaining axioms. But either one is eliminable and provable from the remaining ones, so including both is redundant. The reason for the redundant inclusion is symmetry: There is no basis for choosing among (8a) and (8b) — each is derivable from the other by duality. Choosing just one as an axiom would imply that it was basic and the other "derived", creating an asymmetry where there is none. So here symmetry conflicts with redundancy and symmetry wins.

Lest the reader think Boolean algebra is atypical, here is a more fundamental example. A group is set G with an associative binary relation $\bullet$ satisfying two additional axioms:

(9) a. Identities: There is an $e \in \mathrm{G}$ such that for all $x \in \mathrm{G}$, $e \bullet x = x$ and $x \bullet e = x$.
b. Inverses: For all $x \in \mathrm{G}$ there is a $y \in \mathrm{G}$ such that $y \bullet x = e$ and $x \bullet y = e$.

Now the second conjuncts of (9a) and (9b) can be simultaneously eliminated and proven from the remaining axioms. But doing that implicates that having a left identity element and a left inverse is more basic than having a right identity and a right inverse — their existence being a "mere" theorem, not axiomatic. Again these implicatures introduce an unwarranted asymmetry. We can in fact keep both right hand conjuncts in (9a,b) and eliminate the two left hand ones, deriving them as theorems. Again symmetry trumps redundancy.

4 A Goal of Descriptive Linguistics: Classify Human Grammars by Their Symmetry Groups

How many ways are there to build a Predicate-Argument system? A Modifier system? These questions suggest lower bounds on the expressivity of NLs. We just hint at how to flesh them out. First, for an arbitrary domain E, write Pn for $P(E^n)$, the set of n-ary relations over E (the subsets of En). Write $[A \to B]$ for the set of functions from A into B. Then a minimal Predicate-Argument system is the set of Pn, $0 \leq n \leq 3$ with argument algebras $[Pn + 1 \to Pn], 0 \leq n \leq 2$. A modifer system contains $[Pn \to Pn]$ for $n = 1, 2$ at least, where the functions are restricting: $F(p) \subseteq p$. See Keenan (1981). Obviously we need much more in terms of boundary conditions on denotable objects: quantifiers deriving arguments from $P1$'s, nominalizers of various sorts deriving arguments from Pn's, boolean and binding operators, etc.

References

Aissen, Judith. 1987. *Tzotzil clause structure*. Dordrecht: Reidel.

Baker, Mark C. 2003. *Lexical categories: Verbs, nouns, and adjectives*. NY: Cambridge University Press.

Bar-Hillel, Y., C. Gaifman, and E. Shamir. 1960. On categorial and phrase structure grammars. *Bulletin of the Research Council of Israel* F9:155–166. Reprinted in Y. Bar-Hillel, *Language and Information: Selected Essays on their Theory and Application*. NY: Addison-Wesley, 1964.

Bobaljik, Jonathan David. 2002. A-chains at the interfaces: Copies, agreement and covert movement. *Natural Language and Linguistic Theory* 20:197–267.

Borer, Hagit. 2005a. *In name only*, volume I of *Structuring Sense*. Oxford: Oxford University Press.

Borer, Hagit. 2005b. *The nominal course of events*, volume II of *Structuring Sense*. Oxford: Oxford University Press.

Bošković, Željko, and Howard Lasnik. 2007. *Minimalist syntax: The essential readings*. Oxford: Blackwell.

Brewer, J. W., and M. K. Smith, ed. 1981. *Emmy noether: A tribute to her life and work*. NY: Marcel Dekker.

Bunch, Bryan. 1989. *Reality's mirror*. NY: Wiley.

Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, Massachusetts: MIT Press.

Chomsky, Noam. 1975. *Reflections on language*. NY: Pantheon.

Chomsky, Noam. 1986. *Knowledge of language*. NY: Praeger.

Cole, K. C. 1997. *The universe and the teacup*. NY: Harcourt Brace.

Cole, Peter, and Gabriella Hermon. 2008. VP raising in a VOS language. *Syntax* 11:144–197.

Cotton, Albert F., ed. 1990. *Chemical applications of group theory 3rd edition*. NY: Wiley.

Darvas, G. 2007. *Symmetry*. Boston: Birkhäuser.

Derbyshire, Desmond C. 1977. Word order universals and the existence of ovs languages. *Linguistic Inquiry* 8:590–599.

Gardner, Martin. 2005. *The new ambidextrous universe 3rd revised edition*. NY: Dover.

Gordon, Raymond G., and Barbara F. Grimes, ed. 2005. *Ethnologue: Languages of the world, 15th edition*. Dallas: SIL.

Henkin, Leon. 1961. Some remarks on infinitely long formulas. In *Proceedings of the 1959 Warsaw Symposium on Foundations of Mathematics: Infinitistic Methods*.

Higginbotham, James. 1985. On semantics. *Linguistic Inquiry* 16:547–593.

Hopper, Paul J., and Elizabeth Closs Traugott. 1993. *Grammaticalization*. NY: Cambridge University Press.

Hornstein, Norbert. 1995. *Logical form: From gb to minimalism*. Oxford: Blackwell.

Jakobson, Roman. 1963. Implications of language universals for linguistics. In *Universals of language: report of a conference held at dobbs ferry, new york, april 13-15, 1961*, ed. Joseph Greenberg. Cambridge, Massachusetts: MIT Press.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, Massachusetts: MIT Press.

Keenan, Edward L. 1976. Remarkable subjects in Malagasy. In *Subject and topic*, ed. Charles N. Li. NY: Academic Press.

Keenan, Edward L. 1981. A boolean approach to semantics. In *Formal methods in the study of language*, ed. J. A. G. Groenendijk, T. M. V. Janssen, and M. B. J. Stokhof. Amsterdam: Mathematisch Centrum.

Keenan, Edward L. 1988. On semantics and the binding theory. In *Explaining language universals*, ed. J. Hawkins, 105–145. Oxford: Blackwell.

Keenan, Edward L. 2001. Logical objects. In *Logic, meaning and computation: Essays in memory of alonzo church*, ed. C.A. Anderson and M. Zeleny, 149–180. Boston: Kluwer.

Keenan, Edward L. 2004. Further excursions in natural logic: The mid-point theorems. In *Explaining language universals*, ed. Fritz Hamm and Stephan Kepser. Oxford: Oxford University Press.

Keenan, Edward L., and Edward P. Stabler. 2003. *Bare grammar: Lectures on linguistic invariants*. Stanford, California: CSLI Publications.

Klein, Felix. 1893. A comparative review of recent researches in geometry: Programme on entering the philosophical faculty and the senate of the university of Erlangen in 1872. *Bulletin of the New York Mathematical Society* 2:215–249. Translation by M.W. Haskell of the original October 1872 publication, with a prefatory note by the author.

Kobele, Gregory M. 2006. Generating copies: An investigation into structural identity in language and grammar. Doctoral Dissertation, UCLA.

Lee, Thomas Hun Tak. 1986. Acquisition of quantificational scope in Mandarin Chinese. In *Papers and Reports on Child Language Development, No. 25*. Standford University.

Lewis, David K. 1970. General semantics. *Synthese* 22:18–67. Reprinted in D. Davidson and G. Harman, eds., *Semantics of Natural Language*. Dordrecht: Reidel, 1972.

Montague, Richard. 1973. The proper treatment of quantification in ordinary English. In *Approaches to natural language*, ed. J. Hintikka, J.M.E. Moravcsik, and P. Suppes. Dordrecht: Reidel. Reprinted in R.H. Thomason, editor, *Formal Philosophy: Selected Papers of Richard Montague.* New Haven: Yale University Press, §8.

Radford, Andrew. 1997. *Syntactic theory and the structure of English: A minimalist approach*. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.

Rotman, Joseph. 1999. *An introduction to the theory of groups*. NY: Springer.

Schachter, Paul. 1984. Studies in the structure of Toba Batak. Technical Report UCLA Occasional Papers in Linguistics, Number 5, UCLA, Los Angeles.

Sternberg, S. 1994. *Group theory and physics*. NY: Cambridge University Press.

Stewart, Ian, and Martin Golubitsky. 1999. *Symmetry*. Oxford: Blackwell.

Strocchi, Franco. 2005. *Symmetry breaking*. Lecture Notes in Physics 732. NY: Springer.

Tarski, Alfred. 1986. What are logical notions? *History and Philosophy of Logic* 7:143–154.

van Fraassen, Bas C. 1989. *Laws and symmetry*. Oxford: Clarendon.

Weyl, Hermann. 1952. *Symmetry*. Princeton, New Jersey: Princeton University Press.

Wigner, Eugene P. 1979. *Symmetries and reflections: Scientific essays*. Woodbridge, Connecticut: Oxbow Press.

Zee, A. 1986. *Fearful symmetry*. NY: Macmillan.

Affiliation

Edward L. Keenan
Department of Linguistics
University of California, Los Angeles
keenan@humnet.ucla.edu

Edward P. Stabler
Department of Linguistics
University of California, Los Angeles
stabler@ucla.edu