

Mary-Louise Kean On a Theory of Markedness: Some General Considerations and a Case in Point

0. Introduction

It would generally be agreed that to talk about distinctions between grammatical and ungrammatical strings or to assign some measure of varying degrees of grammaticality to strings would be absurd in the absence of a clear conception of what a theory of grammar is a theory of. By the same token, to stipulate that some rule, string, etc., is marked or unmarked, or more or less marked with respect to some other rule, string, etc., would make little sense in the absence of a coherent conception of what a theory of markedness is a theory of. While, with increasing frequency, one encounters references to markedness in the literature, such discussions as there are typically take place in the absence of any attempt to define what the proper domain of markedness theories is or to characterize their role in general linguistic theory. Rather, what one usually finds is that there is an implied assumption that it is well-understood what theories of markedness are all about. However, consideration of the diversity of uses encountered and the tacit assumptions which they entail strongly suggests that there is little ground for assuming that the domain of markedness is in fact well-understood. In this paper the question to be addressed is then: what is a theory of markedness a theory of, and what is the role of such a theory in grammar? Having proposed an answer to that question, we will then turn to a specific example of a markedness theory, a theory of intrasegmental structure. Having outlined that case we will return again to considerations of markedness theories in general.

1. The Structure of Markedness Theories and the Markedness of Core Grammar.

As a first approximation of an answer to the question with which we are concerned, it is proposed that:

- A. A theory of markedness is a formal theory of substantive universals.
- B. A theory of markedness will assign « probabilities » to the occurrence and variation of substantive elements within and across grammars.
- C. A theory of markedness includes non parameterized constraints which provide the substantive definition of possible grammar of a human language.

From C it follows that

- D. A theory of markedness will set parameterized constraints which set the limits on the substantive variation among languages.

A, above, stipulates the domain of a markedness theory, while *B*, *C*, and *D* set out the function of the markedness theory (at least in part). Taking A-D as a plausible preliminary answer to the central question, the role of markedness in universal grammar (UG) and its contribution to an account of language acquisition become immediately apparent. Taken together with the other components of UG, e.g., a theory of formal universals, A-D will provide an account of the set of simple and abstract core grammars, and also provide an account of the possible departures from core grammar in particular grammars, assigning to such departures their relative « cost ». *A* and *C* define the domain of learning; *A*, *B*, *C* and *D* characterize the extent to which grammars are scattered, and impose constraints, including constraints of plausibility, on the child's hypothesis construction as he approaches the acquisition of the grammar of his language, hypotheses subject, of course, to the constraints of the evaluation procedure (« simplicity metric ») as well.

It is assumed that *B* is not in dispute, that most everyone would agree that a markedness theory of some domain must provide a ranking of the substantive elements of that domain. Therefore, what is at

issue is
marke
of the
marked
the thec
categor
Giv
for any
characte
stated.
expressi
motivati
marked
If a
from con
indisting
gramma
to be a r
find « m
the two r
 $C_i \neq C_j$
cedure (markedn
meet A-L
cumstanc
markedne
that such
is not in s
ed with r
theory of
grammar
the hypoth
tioned V (no more o
ed Adv (o
I can see,
such rules
evaluation
plausible to

issue is whether or not A , C , and hence D are justified. Any theory of markedness which satisfies B will, of necessity, express every element of the domain in question. Thus, just as the theory of segmental markedness is the expression of the set of distinctive features, so too the theory of markedness of (concatenated) categories will express all categories.

Given this, it follows that a theory of markedness satisfying B for any domain will contain all the information relevant to the characterizations of C and D , and therefore be a theory A if principally stated. To treat C and D independently of B would force redundant expression of elements within UG. Since such redundancy is unmotivated it must be the case that we take it *a priori* that a markedness theory must have properties $A-D$.

If a theory of markedness were simply a count of deviations from core grammar and not a theory meeting $A-D$, then it would be indistinguishable from the evaluation procedure over particular grammars as far as rules are concerned. Assuming the rule Move-NP to be a rule of core grammar, where in particular grammars one may find « marked » rules which mention categorial constants, consider the two rules with the factors: (a) $X C_1 Y NP Z$, and (b) $X C_2 Y NP Z$, $C_1 \neq C_2$. These two rules are of equal value under the evaluation procedure (simplicity metric) for any C_1, C_2 , and, therefore, if markedness only counts deviations from core grammar (and does not meet $A-D$) then (a) is as marked as (b) for any C_1, C_2 . Under such circumstances then, such pairs of rules could provide no evidence for a markedness theory. There is, however, no reason to assume *a priori* that such rules should be equally « marked », that their markedness is not in some degree a function of the particular constants mentioned with respect to NP. That is simply to say that by assuming a theory of markedness only to be a count of deviations from core grammar that we would be committing ourselves from the outset to the hypothesis that, for example, a rule of NP movement which mentioned V (or some phrasal projection of V) as a categorial constant is no more or less marked than a rule of NP movement which mentioned Adv (or some phrasal projection) as a constant. There is, as far as I can see, no motivation or justification for such a claim. While all such rules represent deviations from core grammar, and by the evaluation procedure are assigned some non minimal cost, it is plausible to assume that their markedness is in part a function of the

relative markedness of the constant with respect to NP under some theory of markedness of (the concatenation of) categories.

A similar, and perhaps more compelling, line of argument can be made with respect to those rules/structures which involve language particular idiosyncratic suspension (« relaxations ») of the « constraints » e.g., the conditions on binding. Consider the following possibilities: In language A there exists such a suspension with respect to constructions containing some particular (class of) constant(s); in language B there exists a similar suspension involving a different (class of) constant(s); and in language C, the extreme case, there is an across the board suspension. Note, if one allows there to be particular cases of suspensions, then one is, without adding some further stipulation admitting the possibility of language C. Each of these languages will be « marked » by a theory of markedness which is a simple counting procedure of deviations from core grammar. The question is raised however, since in language C there need be no stipulation as to the domains of the suspension then, by a procedure such as the simplicity metric, since no constants would have to be mentioned in the characterization of the domains of suspensions, we would seemingly be forced to conclude that C was in some sense less marked than A on B and that A and B were equivalently marked.

It seems to me that any conception of a theory of markedness which would lead to such conclusions would have to be rejected. Note that simply stipulating that there can be no across the board suspensions while maintaining that the theory of markedness is a simple a count of deviations from core grammar will hardly improve the matter. Under an \bar{X} theory, a suspension involving some occurrences of the categories that are $[+F_i]$ would apparently be less « marked » than those involving suspensions of the categories $[+F_i, +F_j]$. That is, by evaluation procedure type counting we would come to the conclusion that the former language is « less marked » than the latter language even though the former language more closely approaches the prohibited extreme of across the board suspension. Whether, in such cases, it is the former or the latter language which is marked is hardly a topic about which one can have intuitions; it is an empirical question. It is quite conceivable that for some specific F_i the suspension involving the specified constants $[+F_i]$ would be more marked than a suspension involving the specified constant categories $[+F_i, +F_j]$.

In which a tinctive gestures th involves a theorem Now, we core gra recipro: for refle nouns. S suggest tl in the cas mar for t

It m such data of an imj analyses : and impli arbitrary) D, we ha states that ment B » : ment B ». tion betwe C are unre well prove quired, to i mit oneself substantive open-mind commitmen universals a plained by theoretical :

If the l substantive regard then redundancy

In several of the papers presented here analyses are put forward which appear to dictate that a theory of markedness must make distinctive reference to categories. For example, Chomsky (1981) suggests that the appropriate analysis of the reciprocal in English involves a relaxation of the conditions on binding to the special case (a theorem of those conditions) of the Specified Subject Constraint. Now, were a theory of markedness simply a count of deviations from core grammar it would be predicted that as likely as the English reciprocal case would be the case where the conditions were relaxed for reflexives, or the conditions were relaxed for nonreflexive pronouns. Surely that is implausible; Lasnik and Freidin (1981) in fact suggest that one only finds a relaxation of the conditions on binding in the case of reflexives where there is the same relaxation in a grammar for reciprocals.

It might be suggested that the appropriate way to account for such data is not within a theory of markedness, but rather by means of an implicational universal or a redundancy rule. However, such analyses are quite unsatisfying from the outset. Redundancy rules and implicational universals have the property of stating apparently arbitrary relations between elements within a domain. In the domain D , we have an implicational universal (or redundancy rule) which states that « if there exists an element A then there also exists an element B » and « if there exists an element C then there exists an element B ». While A and C both imply B , nothing is said of the relation between C and A : it might equally well be the case that (a) A and C are unrelated, (b) if A then C , or (c) if C then A . While it might well prove that implicational universals and redundancy rules are required, to make appeal to such constructs from the outset is to commit oneself to the position that the internal structure of the set of substantive elements of any domain is completely arbitrary. A more open-minded view would certainly be to make no such theoretical commitment, and, instead, assume that so-called implicational universals and redundancy rules are in fact data which are to be explained by the theory of grammar and are not themselves a theoretical account.

If the lessons of phonology where the most extensive work on substantive elements has been carried out hold any moral in this regard then it is to be expected that implicational universals and redundancy rules will most likely have no status in the theory of any

domain of substantive elements. Having been proposed and utilized in phonology, with the development of theories of markedness to account for some properties of segmental alternation it became evident that those theories also provided an account for the data of implicational universals and redundancy rules as characterizations of the structure of segmental systems (Kean 1975, van Lessen Kloeke, 1979, 1981). There is *prima facie* evidence that the same is to be expected in other substantive domains. One such example will be briefly considered.

In languages with morphologically overt tense marking on verbs, the tenses are [\pm past] or [\pm future]. In such languages if there is no morphologically specified subjunctive, tense plays a crucial non semantically temporal role in the statement of counterfactual conditionals: the nonpresent (i.e., [+ past] or [+ future]) is interpreted as negation in the present. Thus, in English if I say *If I were the queen of Sheba...* that means that I am not now the queen of Sheba (though I might have been in the past); if it were the case that I am now the queen of Sheba (though I might not have been in the past) and wanted to refer in such a conditional to some previous point in time when I wasn't the queen of Sheba, then the perfective would be the syntactic marker which received the interpretation of not present, i.e., *If I had been the queen of Sheba...* means that there is sometime in the past when I wasn't the queen of Sheba (though I might well be the queen of Sheba now). Not all languages have overt morphological tense. In those languages without overt tense marking (e.g. Klamath) one apparently invariably has an overt morphological distinction between indicatives and subjunctives. Based on these data we can postulate a series of implicational universals:

- a. If a language has a morphological subjunctive then it also has a morphological indicative.
- b. If a language has no morphological tense then that language has a morphological indicative/subjunctive distinction.

The first of these implications is less than illuminating; if all languages have indicatives, then given the possibility of a morphological subjunctive it follows as a matter of logical necessity that the language will have a morphological indicative (even if the indicative has a phonetically empty realization). The second implica-

tion seem:
there is s
however,
say about
marking f
junctive ir
might, giv
on having
respect to
which ther
junctive. F
ing no mor
the subjunc
conditions
tense carrie
universals a
terconnectic
evaluation c
would, how
arbitrarily e

It seems
of markedn
categories of
fashion. Unl
simple count
mind, being
markedness a
the possibilit
important are
sal grammar.
markedness s
grammar is eq
horse in.

A second
edness in the
mar is the « ur
this notion: on
that is, is liter
that core gram

tion seems somewhat more interesting in that it captures the fact that there is some apparent relation between tense and mood. It too, however, is not particularly revealing. What, for example, are we to say about languages like Spanish where there is overt morphological marking for both tense and mood? Should we conclude that the subjunctive in Spanish is just some costly embellishment? While one might, given more data, want to argue that there is a cost attendant on having a subjunctive in a language with morphological tense with respect to counterfactuals, presumably there will be other respects in which there is a cost attendant on not having a morphological subjunctive. Furthermore, one must ask what is the relative cost of having no morphological tense in the first place; surely in such languages the subjunctive will be relatively less costly, but there must be other conditions in such languages where the absence of morphological tense carries with it some cost. By their very nature implicational universals and redundancy rules are incapable of expressing such interconnections; any theory which did express and provide for an evaluation of the tension between tense marking and mood marking would, however, of necessity contain just that information which is arbitrarily expressed by the implicational universals (a) and (b).

It seems to me necessary to assume from the outset that a theory of markedness must make reference to the specific substantive categories of the domain, arrayed hierarchically in some principled fashion. Unless this is done we are left with the evaluation procedure, simple counting, *ad hoc* stipulation, etc. Even worse, at least to my mind, being content with simply characterizing a theory of markedness as a counting of deviations from core grammar cuts off the possibility of inquiry into what is a most interesting and surely important area of study for the development of the theory of universal grammar. To pervert a popular cliché, taking the position that markedness simply involves the counting of deviations from core grammar is equivalent to shutting the barn door before you've let the horse in.

A second problem I find with many characterizations of markedness in the literature is an ambiguity in the claim that core grammar is the « unmarked » case. There are two ways we might interpret this notion: on the one hand, that core grammar has no marks at all, that is, is literally and completely unmarked, or, on the other hand that core grammar represents the least marked possible case, that it is

less marked than any other case.¹ Note, if we take it that a theory of markedness simply measures deviations from core grammar, then we are committing ourselves to the former interpretation, while if we take the view of markedness *A-D* we are only committing ourselves to the latter position, leaving it open as an empirical question whether or not core grammar is completely unmarked. Just as a matter of considering what would be the potentially most fruitful assumption to take in terms of any attempt to develop an explicit theory of markedness, clearly the latter possibility would have to be selected. It is, however, not just a matter of desirable ways to pursue research which mitigates against the former view. If core grammar is literally unmarked, then markedness theory will not include any of the principles of core grammar. In that case, *C* and perhaps part of *B*, would not be components of the markedness theory and the representation of the substantive domain in question, *A*, would have to be stated twice, once for core grammar along with *C* (and *D*) and once for the markedness theory along with *B*. That is, we would be inviting redundancy without motivation from the outset. Secondly, if government and case relations are to be stated in terms of a markedness theory as van Riemsdijk suggests, then it is clear that the assumption that core grammar is literally unmarked cannot be maintained unless one is prepared to argue that *oblique* is not a case in « core grammar ». One can extend this line of argument to absurdity where one would be forced to conclude that core grammar, the unmarked case, is in fact a grammar defined substantively on the empty set.

That core grammar is not literally unmarked, only relatively unmarked, can be demonstrated quite clearly, I think, by consideration of the structure of segmental systems. There is an implied tendency among many linguists to think that core grammar is a notion which ranges only over syntax and logical form, a view I find to be as unsupported as the notion that there is not a general definition of

¹ Rouveret and Vergnaud (1980) propose that markedness is a measure of deviations from core grammars through the successive sequence of grammars the child entertains in the course of acquisition. Each successive grammar is more marked than the previous grammar, the grammar G_n of the final state being the most highly marked. Their conception of a markedness theory is based on the assumption that core grammar be taken to be « the stock of hypotheses of null cost » (p. 195). The Rouveret and Vergnaud proposal is discussed in Section 3.3.

autonom
(lexical)
by all acc
must be :

In th
/a/; in m
segments
marked c
sist of at
be no m
one consi
a system
must be c
that just
have the
language
tion that
require a
theoretic
segments
what we
segments
structure
theory of
account.
tories in
if that ec
then it ca
There is ;
tion, ind

If w
also mak
try to ma
but one c
grammar
of any s
would be
does not
could not

autonomy which extends beyond the relation between syntax and (lexical) semantics. As phonology is a component of core grammar by all accounts (Chomsky and Lasnik, 1977; Chomsky, 1980), there must be an « unmarked » phonology.

In the SPE markedness theory there is one unmarked segment, /a/; in my markedness theory (Kean, 1975) there are two unmarked segments /t/ and /a/. If we take core grammar to be the literally unmarked case, then the segmental system of core grammar would consist of at most two segments. Note, such a segmental inventory would be no more or less marked than one consisting of one segment, or one consisting of no segments. Since to be a natural human language, a system must have a segmental inventory, then that is a fact which must be captured in a theory of core grammar. Now, it seems to me that just as it is transparently obvious that a human language cannot have the empty set as its segmental inventory, so too a natural language cannot have just one or two segments. To maintain the notion that core grammar is literally unmarked in the face of this would require appeal to, for example, some non-grammatical information theoretic account to explain why languages never could have just two segments; given such an account it would have to be the case that what we take to be part of the grammatical theory of markedness of segments — that part which addresses the markedness of the internal structure of sound systems — would not be a component of the theory of grammar at all. Surely this would hardly be a satisfactory account. It would seem then that any account of segmental inventories in core grammar would have to include marked segments, and if that component of core grammar does include marked segments then it cannot be the case that core grammar is literally unmarked. There is an inevitable cost associated with any substantive specification, independent of domain.

If we take core grammar to be literally unmarked then we are also making the claim that there is but one core grammar. One might try to maintain that core grammar is literally unmarked and there is but one core grammar by claiming that core grammar is a schema for grammars and makes no substantive distinctions among the elements of any substantive domain. Thus, for example, Move-Category would be the transformation of core grammar, and since that rule does not distinguish among categories (major or minor) then one could not associate any marking or cost with it. To take such a view

t theory of
ur, then we
while if we
g ourselves
d question
st as a mat-
st fruitful
an explicit
have to be
s to pursue
grammar is
lude any of
aps part of
ry and the
would have
(and *D*) and
ve would be
Secondly, if
terms of a
lear that the
not be main-
not a case in
to absurdity
mar, the un-
on the empty

relatively un-
consideration
lied tendency
notion which
l to be as un-
definition of

measure of devia-
mmars the child
ore marked than
ost highly mark-
mption that core
» (p. 195). The

would be to reduce the theory of core grammar to a theory of formal universals. Such a reduction is incompatible with a significant amount of the research which has been done on core grammar. For example, it is not a position which can be maintained at the same time as the claim that there is a rule of reciprocal interpretation in core grammar (and every grammar must have a distinctive expression of the reciprocal).

To assume that there is just one grammar which is the core grammar for all languages, and not that there is a set of grammars which are the core grammars is, it seems to me, an untenable position. Consider again the case of tense and mood. Those are two parameters along which languages can vary, with some cost attendant on their variation. Any theory of core grammar will, it is assumed, include schema for the interpretation of time as it is syntactically represented; this being so the grammatical features of tense and mood must be available to the theory of grammar which expresses core grammar. Given that there are costs associated with fixing those parameters in one way or another, it follows that core grammar cannot be literally unmarked. While it is impossible to predict whether or not there will be some combination of tense and mood marking which is less marked, less costly, than all others, there is, it should be clear, no reason to anticipate such an eventuality. Taking together all the substantive parameters (case, tense, mood, plural/distributive, etc.) along which languages can vary with their contingent relations and associated cost, that there should be a unique least marked case seems quite implausible. Rather what is suggested is that there will be a finite set of relatively scattered grammars which are core grammars, core grammars being those grammars which involve no relaxation of conditions and constraints, where rules have minimal expressive power, and where the set of substantive elements expressed in the grammar are minimally marked. In fact, phonology provides evidence that this is so. The markedness theory of segmental structure to which we will turn directly specifies a set of minimally marked segmental inventories, not a unique minimally unmarked segmental inventory.

2. A Case

2.1. The

Having m:
markedne

There
markedne:
ories of n
which mig
former to
(1925); th
to the sup
and propo
have invar
is, segmen
pretation
subsystem
and the in
tions, but
dencies. O
there is litt

¹ As m
theory in ph
features. A s
from the stan
I do not take
sal Phonetic
pressed in the
phonetically
mapping of t
biguity of the
Universal Ph
component o
their so-called
the acoustic p
twelve memb
tative as the
arguments fo

Basbøll
ed feature sys
While the fe
underlying th
and Halle.

2. A Case in Point

2.1. *The Domain of Phonological Markedness*

Having made these rather general remarks, let me turn to a particular markedness theory, the markedness theory of phonology.

There are two issues which must be addressed by any theory of markedness in phonology: (a) the structure of the segmental inventories of natural language, and (b) a set of problems relating to rules which might be termed 'failures of the evaluation procedure'.² The former topic has not been systematically addressed since Sapir (1925); the best one finds in the literature are stipulations, allusions to the supposed symmetry of segmental systems, redundancy rules, and proposals for the so-called implicational universals, all of which have invariably been inconclusive and ultimately misdirected. That is, segmental systems are not symmetrical under any obvious interpretation of the notion symmetry (though it is, of course, true that subsystems in segmental inventories are often sort of symmetrical), and the implicational universals aren't absolute universal implications, but rather usually only characterizations of some strong tendencies. On the matter of implicational universals, as was noted above, there is little reason to believe that they have any systematic status in

² As markedness theories are theories of substantive elements, the markedness theory in phonology will be some sort of theory of segments, or, more precisely, features. A segment is taken to be a set of specified distinctive features. Departing from the standard theory (cf., Chomsky 1965; Chomsky and Halle 1968, Halle 1959), I do not take the set of distinctive features in phonology to be expressed within Universal Phonetic Theory (Kean 1978). I assume that the set of phonological features is expressed in the markedness theory, and that the features so expressed are abstract and phonetically ambiguous. I assume that Universal Phonetic Theory is a theory of the mapping of these features onto their possible phonetic realizations, and that the ambiguity of the phonological features is vitiated by the constraints which are imposed by Universal Phonetic Theory on that mapping. (Appendix II gives a proposal for that component of Universal Phonetic Theory which maps phonological features onto their so-called articulatory phonetic instantiations; it is left here an open question what the acoustic phonetic analogue is.) I assume that the set of phonological features has twelve members (see (10) below); this assumption should, however, be viewed as tentative as there is still considerable detail to be worked out. In section 2.3 some arguments for a small feature system are given.

Basbøll and Bierwisch have observed (personal communication) that this assumed feature system is reminiscent of that proposed in Jakobson, Fant, and Halle (1952). While the features of (10) are not acoustic features, many of the considerations underlying the motivation of such a system are the same as those of Jakobson, Fant, and Halle.

the theory of grammar; to the extent that any apparently hold, it is, at least in phonology and probably in other domains as well, in virtue of their being derivative from markedness theory. The latter set of problems — those relating to rules — was first noted by Chomsky and Halle (1968), who made the following observations.

- (1) *Alternation is not symmetrical*
k-č is a relatively frequent rule, whereas one doesn't seem to find č → k.
- (2) *Some rules are « too expensive »*
k-č takes four features to state (SPE feature system) (N.B. voicing alternation takes one feature to state, and spirantization takes one or two features.)

There are some additional problems of alternation for the standard theory of phonology which Chomsky and Halle did not observe.

- (3) *The apparent non-occurrence of some maximally simple alternations*
t- [+son] is an apparently unattested rule
- (4) *Phonological rules do not compound*
One finds the rules [-son] → [-back] and [-son] → [+cont], but apparently not [-son] → [-back] [+cont]
- (5) *Ad hoc alternations*
Under lenition, while *p* becomes *β* and *k* becomes *γ*, *t*, curiously, does not typically become *δ* but rather becomes *r*.

Yet another problem with rules is the question of the relation between the features specified by a rule and the triggering environment; Dresher (1979) outlines some of the central features of this area. I am assuming that the answer to the question of the markedness of strings (Cairns 1969) is the same as that to the question of environments. The issue of the markedness of segmental sequences in rules and in strings will not be considered here.

2.2. An Intra-segmental Markedness Theory - A, B

Having made these preliminary remarks sketching out the issues with which the markedness theory for phonology will be concerned, I want to turn now to outlining a theory of markedness which attempts to address each of these problems.

A-I

A-II

Given A-I are
is, intrinsic
ations:

(6)

F
F
F

Let Φ be the set of phonological features

$$\Phi = \{F_1, \dots, F_n\}$$

A-I For each feature F_i there exists one and only one *markedness convention* of the form:

$$a \quad uF_i, \beta_1 G_1, \dots, \beta_m G_m - \alpha F_i, \beta_1 G_1, \dots, \beta_m G_m$$

where: $\alpha, \beta_1, \dots, \beta_m$ are + or —,
and F_i, G_1, \dots, G_m
are features

which is understood as

a' The unmarked specification of the feature F_i is rewritten as + or — in any segment which is $\beta_1 G_1, \dots, \beta_m G_m, \beta_1, \dots, \beta_m, +$ or —, and G_1, \dots, G_m , features.

Using standard phonological notations,

$$a'' \quad [u F_i] - [\alpha F_i] / X$$

where: X is

$$\begin{bmatrix} \overline{\beta_1 G_1} \\ \vdots \\ \beta_m G_m \end{bmatrix}$$

$\alpha, \beta_1, \dots, \beta_m$ are + or —, and
 F_i, G_1, \dots, G_m are features

A-II For any pair of features $F_i, F_j, F_i \neq F_j$, if F_i is a specified feature in the markedness convention of the feature F_j , then F_j is not a specified feature in the markedness convention of F_i .

Given A-I and A-II, it follows that the set of markedness conventions is, intrinsically, a partially ordered set, admitting the ordering relations:

$$(6) \quad \begin{array}{c} F_i \\ F_j \\ F_k \end{array} \quad \begin{array}{c} F_i \\ / \quad \backslash \\ F_j \quad F_k \end{array} \quad \begin{array}{c} F_i \quad F_j \\ \backslash \quad / \\ F_k \end{array}$$

where: F_i, F_j, F_k stand for the markedness conventions for the features F_i, F_j, F_k , respectively.

It should be evident that unless X, the environment of the markedness convention, is null, the convention will only characterize the markedness of the specification of a feature in a proper subset of the set of segments. Since what is at issue is the markedness of the specification of any feature in any segment, an interpretive convention is required. The convention proposed is the Complement Convention, a convention by which every markedness convention is interpreted as a set of *markedness rules*.³

A-III. Complement Convention

$[u F_i] - [\alpha F_i] / X (= A-I)$

$[m F_i] - [-\alpha F_i] / X$

$[u F_i] - [-\alpha F_i] / \bar{X}$

$[m F_i] - [\alpha F_i] / \bar{X}$

The markedness theory characterized by A-I, A-II, and A-III is a formal theory of the substantive universals of phonology, the features. The theory then satisfies the descriptive criterion A put forward above.

Using features with which most people should be familiar, consider the feature *low* from the SPE system. It is « unmarked » for consonants and all vowels save *a* to be [$- \text{low}$]. To capture this it might be proposed that the markedness convention for *low* is:

$$(7) \quad [u \text{ low}] - [+ \text{ low}] / \begin{bmatrix} - \text{cons} \\ + \text{back} \end{bmatrix}$$

Accepting (7) as our markedness convention for *low* means by A-II that the conventions for *consonantal* and *back* must be ordered before that for *low* and they may not refer to the feature *low*. Inter-

³ For a discussion comparing the markedness theory A-I, II, III with that in Chomsky and Halle (1968), see Kean (1975). The essential differences between the two theories are: (a) In SPE there are no particular formal constraints on the form of markedness conventions, their expressive power being unlimited up to the expressive power of phonological rules given a set of features and four specifications (u, m, +, -); (b) The markedness conventions for all features (save *consonantal* and *vocalic*) which apply in e.g., vowels, bear no relation to the markedness conventions for consonants, liquids and glides; for each of these four classes there is an autonomous set of markedness conventions so that the SPE theory entails the claim that there are no general principles which cross-classify in markedness theory the set of members of these sets (except with respect to *vocalic* and *consonantal*). (c) The markedness conventions of the SPE system are extrinsically ordered with different orderings obtaining for the different classes of segments.

preting
marked

(8)

By case (i)
low. By (ii)
are marked
the nonlow
case (8d)
e.g., *ae*, *e*

⁴ A word
ple is an un
clear that v
marked in s
segment S
take it that
marked with
be marked v
unmarked f
which says t
this case w
necessary th
set character
its simplicity
thoroughly r
not go throu
germane to c
sion of the r

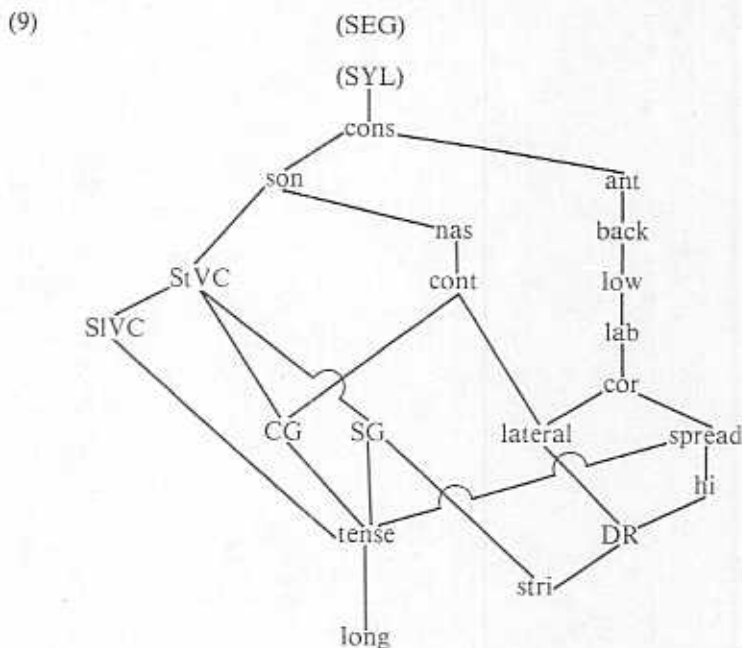
preting (7) by the Complement Convention A-III we would get the markedness rules in (8).

- (8)
- a. $[u \text{ low}] \rightarrow [+ \text{ low}] / \left[\begin{array}{c} - \text{ cons} \\ + \text{ back} \end{array} \right]$
- b. $[m \text{ low}] \rightarrow [- \text{ low}] / \left[\begin{array}{c} - \text{ cons} \\ + \text{ back} \end{array} \right]$
- c. $[u \text{ low}] \rightarrow [- \text{ low}] / \left\{ \begin{array}{c} [+ \text{ cons}] \\ [- \text{ back}] \end{array} \right\}$
- d. $[m \text{ low}] \rightarrow [+ \text{ low}] / \left\{ \begin{array}{c} [+ \text{ cons}] \\ [- \text{ back}] \end{array} \right\}$

By case (8a) the vowel *a* and the laryngeals *h* and *ʔ* are unmarked for *low*. By (8b) the nonlow back vowels, e.g., *u*, *o*, *ə*, and the glide *w* are marked for *low*. By case (8c) the front vowels, e.g., *i* and *e*, and the nonlow consonants, e.g., *p*, *s*, and *k* are unmarked for *low*. By case (8d) the low consonants, e.g., the stop *g* and low front vowels, e.g., *ae*, are marked for *low*.⁴

⁴ A word of caution is perhaps in order here. Central to constructing this example is an unstated appeal to the intuition that *a* is the unmarked vowel. It should be clear that when one turns to considerations of segments which one assumes to be marked in some degree there can be no intuitions appealed to, that is, to know that segment *S* is marked does not answer the question of how it is marked. Note, if we take it that *a* is the unmarked vowel, it does not follow necessarily that *u* and *o* will be marked with respect to the feature *low*; for all one can antecedently know, they might be marked with respect to the feature *round* by some convention which says that it is unmarked for segments to be [- round], and unmarked for *low* by a convention which says that it is unmarked for rounded segments to be [- low]. Furthermore, in this case while the environment of the convention includes *a*, it is by no means necessary that the convention had to have *the* unmarked segment as a member of the set characterized in the environment. What makes the example work as an example is its simplicity and transparency; however, in having these two properties it is a thoroughly misleading example of what markedness conventions are often like. I will not go through the markedness conventions for all the features as that is not directly germane to our interests here. See van Lessen Kloetze (1979, 1981) for detailed discussion of the motivation of particular markedness conventions.

As noted above, by A-I and A-II, it follows that the set of markedness conventions is intrinsically partially ordered. Given such an ordering of the conventions, it is possible to derive an ordering of the features themselves; if the convention for some feature F_i is intrinsically ordered before the convention for the feature F_j , then we can say the feature F_i is ordered before the feature F_j . Using SPE type features, in my earlier work on markedness (Kean 1975, 1977) the hierarchy of features was that given in (9); in my current work, the hierarchy is essentially a compressed version of (9) as is illustrated in (10). The parenthesized features in (10) are *hints* as to the interpretation of these features in an SPE type feature system.⁵ [Appendix I provides the interpretations for the feature abbreviations used.]



⁵ (10) should be understood as preliminary since the set of phonetic correspondence relations has yet to be fully worked out. The features in (10) receive their phonetic interpretations — in terms of features such as those in parentheses in (10) — by the phonetic theory outlined in Appendix II.

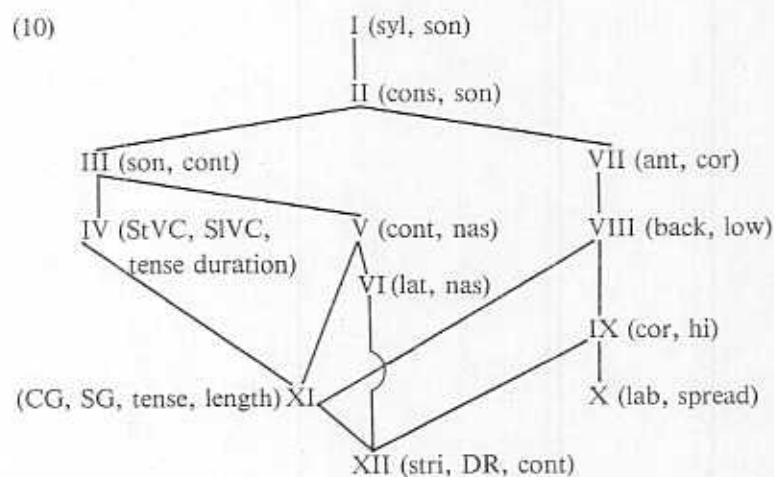
(10)

III
IV

(CG, SG,

As yet, the n
A; it does no
vide any acc
the apparent
that everyone
to have the p
which satisfie
redundantly t
that line of ar
ed in A-I to A
phonology.

Given a r
a set of mar
feature of any
specification,
feature in its p
for every fe
markedness t
relative likelih
features in se
greater the nu
of a segment
highly valued
dancy rules m



As yet, the markedness theory presented here only satisfies criterion *A*; it does not satisfy criteria *B-D*, and, consequently, does not provide any account for the internal structure of segmental systems or the apparent failures of the evaluation procedure. On the assumption that everyone would probably agree that any markedness theory had to have the property *B*, I made the *a priori* argument that a theory which satisfied *B* would be a theory which would also provide non redundantly the information required to satisfy *C* and *D*. Following that line of argument, let us first consider whether the theory outlined in A-I to A-III provides a means for meeting *B* for intrasegmental phonology.

Given a markedness convention for every feature, interpreted as a set of markedness rules by the Complement Convention, every feature of any segment may be represented in terms of its markedness specification, *u* or *m*. Since every segment must be specified for every feature in its phonetic representation, every segment may be specified for every feature in its phonological representation. In non markedness theories of segmental structure, to account for the relative likelihood of occurrence of various specifications of various features in segments, appeal was made to redundancy rules: the greater the number of features which could be redundantly specified of a segment within the segmental system of a language, the more highly valued the segment, *ceteris paribus*. The notion of such redundancy rules makes no sense under a markedness theory such as that

et of mark-
en such an
ring of the
 F_i is intrin-
hen we can
g SPE type
, 1977) the
t work, the
lustrated in
: interpreta-
Appendix I
s used.]

ht
ck
w
b
or
spread
hi
R

of phonetic cor-
(10) receive their
atheses in (10) —

which has been outlined here. Making the most extreme claim, as is suggested above, I would be prepared to argue that there are in fact no redundancy rules in phonology and probably none in grammar, only markedness conventions and rules. Given that redundancy rules make no sense within this system, it is assumed that every segment is specified for every feature at every level representation. Any segment assigned +/- feature specifications can, by the markedness rules, be interpreted directly as a segment whose features are assigned *m/u* specifications. To order the set of segments and segmental systems in terms of their « absolute » complexity, B-I is proposed.

- B-I
- a. The complexity of a segment is equal to the sum of its marked features.
The greater the complexity of a segment by B-I, the less probable it is that that segment will occur in the segmental system of a language at any systematic level of segmental representation.
 - b. The complexity of a segmental system is a function of the complexity of its members.

There is a second set of phenomena that need to be accounted for to satisfy B, the so-called implicational universals. It is clear that there are strong tendencies in languages to meet the implicational universal in segmental systems; languages with *b* typically have *p*, a language with *q* typically has *k*, etc. To account for the data reflected in the implicational universals B-II is proposed; B-II (a) is the general implicational universal of the markedness theory and B-II (b) is its evaluative interpretation.

- B-II
- a. If in the segmental system of a grammar there exists a segment *S* which is marked for some feature F_1 , then that system also contains a segment *S'* which is unmarked for F_1 , the markedness specifications of all other features being equal.
A grammar may « violate » (a), but no language may flaunt it.
 - b. There exists a probability *n*, where *n* is small, assigned to its being the case that the antecedent of (a) is true and the consequent false.

As combined probabilities are multiplicative, not additive, it follows

that, *n*
language

2.3 The

Let us n
tains the
the noti
of the in
marked
on the s
phonolo
cessary t
the ansv
tem emp

The
ly, to c
language
number
(Note, B
likely to
takes the
tal syste
member
proper s
language
to provi
necessary
marked :
stipulatic
to do tha
involved:
« top » c
to find sc
generaliz

⁶ As i
markedness
markedness
dependent o

that, $n < 1$, the number of « violations » of (a) to be found in any language is going to be quite small.

2.3 *The Structure of Segmental Inventories - C*

Let us now turn to the question of whether the system outlined contains the information relevant to satisfying *C*, that is, can we define the notion « possible language » for segmental systems on the basis of the information represented in A-I to A-III, alone? In meeting *C* a markedness theory will have to account for the absolute constraints on the structure of segmental systems, and also the notion « possible phonological alternation ». In addressing the former issue, it is necessary to consider both the feature system of (9) and that of (10) since the answer one provides rests in large measure on the feature system employed.⁶

The question of possible segmental systems amounts, essentially, to consideration of the following questions: why is there no language with the segmental system /*t, a*/ or some other small number of fairly unmarked segments where B-II (a) is not flaunted. (Note, B-II provides (at least a partial) account for why we won't be likely to find a language with the segmental system /*a, λ, q*/.) If one takes the feature system of (9) and considers a wide range of segmental systems, there exists a set of segments where it is true of every member of the set that B-II (a) is (fairly) strictly adhered to over a proper subset of the features, the same proper subset in every language. Presumably that is not just a happy coincidence. In order to provide some sort of principled account for this fact, it is necessary to define the set of features for which there is invariably a marked segment in a language. One could, of course, do this by stipulation, and nothing would be *a priori* wrong with that; however, to do that would be to miss a generalization about the set of features involved: those features which are invariably marked fall at the « top » of the hierarchy. Noticing that, one is then inclined to want to find some principled characterization of the set, and then state the generalization about the invariability of there being marked segments

⁶ As is argued here and in Kean (1975) and in van Lessen Kloeke (1979, 1981), a markedness theory imposes constraints on possible segmental systems. Below markedness arguments for a system such as (10) will be given; those arguments are independent of the arguments for such a feature system given in Kean (1978).

in terms of that principled set. If, however, one attempts to do that, one quickly finds that the « principles » which define the set are absurdly arbitrary, really no better than simply stipulating the members of the set. That this is so is, I think, adequately demonstrated by (11) which almost, but not quite, works.

Definitions: (a) The feature F_i *dominates* the feature F_j ($F_i > F_j$) just in case the convention for F_i is intrinsically ordered before the convention for F_j .

(b) The feature F_i *immediately dominates* the feature F_j just in case (i) $F_i > F_j$, and (ii) there exists no feature F_k such that $F_i > F_k > F_j$.

(11) Definitions: (a) The feature F_i is *optional* if (i) there exist the features $F_j, F_k, F_j \neq F_k$, such that $F_i > F_j$ and $F_k > F_i$, and $F_j \succ F_k$ and $F_k \succ F_j$, and (ii) there exists no $F_l, F_m, F_l \neq F_m$, such that F_i directly dominates F_l and F_m directly dominates F_l .

(b) Any feature F_i is dominated by an optional feature F_j (as defined in (11a)) is itself optional.

Conditions: In every underlying segmental system:

(a) for each feature F_i, F_i *not* optional, there exists at least one segment which is $[m F_i]$, and

(b) for each pair of features F_i, F_j, F_i and F_j , not optional, $F_i \neq F_j$, if there exists a segment S which is marked for both F_i and F_j , then there exists a segment S' such that S' is marked for one but not both F_i and F_j , and S and S' agree in their marking for all other non-optional features.

The effect of the definitions is to characterize the set of features *consonantal, sonorant, nasal, anterior, back, low, and labial* as not optional features. As is implied by the expression « almost, but not

quite, v
capricious
belong
without
can do,
does no

On
there ar
reason t
be the
segment
one wou
segment
ventorie
tories? Y
question
simply
phonetic
which is,
tive is of
leads to
with the
is that th
for each
feature,
attempt t
one can
systems c

C-I

C-I d
tems. It p

quite, works, » there is more reason to be wary of (11) than just the capriciousness of its definitions. It turns out that *nasal* does not belong to the set of non-optional features, i.e., there are languages without segments marked for *nasal*, e.g., Wichita. So the best one can do, it would seem, is to make recourse to stipulation. Stipulation does not, however, solve all the problems.

One problem that arises is that given the set of features in (9) there are only *eight* non-optional (i.e., obligatory) features. The reason this is a problem is that no matter where one looks it seems to be the case that languages never, curiously, have fewer than 13 segments. If there were only eight obligatory features then given (11) one would have to assume that there could be languages with just nine segments. It is easy enough to find thirteen and fourteen segment inventories, so why can't one readily find nine to ten segment inventories? Yet another bothersome oddity with the (9/11) system is the question of why there should be « optional » features at all; are they simply potential linguistic embellishments? Note also that the phonetic phenomena called *voicing* are optional, whereas *anterior* which is, at least in terms of +/— specifications, often non distinctive is obligatory. It is consideration of these sorts of issues which leads to one inevitable conclusion: there must be something wrong with the features themselves. Obviously, the simplest thing to assume is that there are only obligatory features, i.e., that in every language for each feature there exists one segment which is marked for that feature, subject to some conditions (e.g., B-II). (10) represents an attempt to develop such a feature system, and given such a system, one can state the condition on the necessary richness of segmental systems quite generally, as in C-I.

C-I In every grammar

- (a) for every feature F_i there exists at least one segment which is $[m F_i]$, and
- (b) for each pair of features $F_i, F_j, F_i \neq F_j$, if there exists a segment S which is marked for both F_i and F_j , then there exists a segment S' which is marked for one but not both F_i and F_j , all other markedness specification being equal.

C-I does more than set a lower bound on possible segmental systems. It provides an immediate account for why it is that languages

with only two vowels have lots of consonants, frequently fairly exotic consonants. The situation is essentially this: the simplest segments which one can have in a language which allow for the satisfaction of C-I are, for some features, vowels; not having vowels, in order to satisfy C-I one must, in effect, go around Robin Hood's barn. Thus in two vowel languages such as Kabardian and Kaitij, while in absolute terms the segmental inventories are quite marked (B-I), relative to their having two vowels only, they are not so surprisingly or excessively marked with respect to C-I.⁷

The set of features such as that suggested in (10) is not only motivated by markedness consideration of the structure of segmental systems. There are three independent motivations for it. First, as discussed in Kean (1978), a feature system such as that in (9) is consistent with there existing a language with three distinctive *t*'s and three distinctive *d*'s, where the *t*'s vary from each other only in the way in which voicelessness was phonetically instantiated. One does not encounter such languages; rather, what one finds, is, apparently, at most a three way distinction in some languages. This is what is predicted by the feature system such as (10) interpreted phonetically by a Universal Phonetic Theory of the sort given in Appendix II. The phonetic features of voicing stand in the extensions of two features, IV and XI. One principle of the phonetic theory, Disjointness, limits any possible realization to being the realization for just one phonological feature. Thus, the feature *tense* may in any language be the phonetically distinctive realization of IV or XI, but not both. Other principles (exhaustiveness, distinctiveness, and nesting) conspire to limit the options still left open under Disjointness.

The second independent motivation for a small abstract feature system (and an elaborated Universal Phonetic Theory) comes from consideration of the spontaneous segmental substitution errors made by the normal and brain damaged individuals. Under an SPE-type feature system, all phonological features which are distinctive in a language are phonetically directly accessible. Thus, for example, in English since *lateral* and *DR* are distinctive, they are crucially exploited by the language. Now, if we assume that the feature system of

⁷ See Kean (1975) for examples of how a condition such as C-I does provide an account for both large and small segmental systems.

a langua
compon
the case
substitu
stay with
segment
feature s
ticator
feature s
observat
one cou
recruitm
vision (F
course, r

And
provides
and acou
an accou
language
the phon
that is a
phonolog
quisition
represent
constrain
the mark
tion on
phonolog
disparitie

Ther
system: fi
account f
such a fi
segments;
language
fourth, it
which exi

a language is « psychologically real » in the sense of being a potent component of on-line processing in performance, then it should be the case that an English speaker might make the speech error of substituting λ for l or r , etc. But the fact of the matter is that people stay within the segmental systems of their languages when they make segmental substitutions. This observation can only be explained in a feature system which is directly phonetic on the basis of appeal to articulatory habits, which is hardly a satisfying account. Under the feature system here, with a rich theory of phonetic realization, this observation follows from the most obvious sort of learning theory one could associate with the Universal Phonetic Theory, i.e., a recruitment learning theory analogous to that which is assumed for vision (Hubel and Wiesel 1970). Such segmental learning would, of course, not be motor.

Another problem with a feature system such as that in (9) is that it provides no account for the disparities between articulatory features and acoustic phonetic features. There is no mechanism for providing an account for how, given an acoustic signal the child acquiring his language, translates that into the essentially articulatory features of the phonology (9). Under a feature system of the sort I am suggesting that is accounted for by Universal Phonetic Theory. Given the phonological ambiguity of acoustic stimuli, the child's task in acquisition is to determine the correct mapping from the acoustic representation onto the phonological representation, subject to strict constraints — both those of Universal Phonetic Theory and those of the markedness theory (e.g., C-I) — and then from that representation onto the articulatory realization representation. The phonological features then serve the phonetic end of mediating the disparities which exist between the modalities.

There are, then, four motivations for a small abstract feature system: first, given such a feature system one can provide a general account for the internal structure of a segmental system; second, such a feature system does not overgenerate certain classes of segments; third, the system provides a means for accounting for the language bound nature of phonological/phonetic performance; and fourth, it acknowledges and attempts to account for the disparities which exist between perception and production.

2.4 Intrasegmental Markedness and Possible Rules - C

The second area of intrasegmental phonology which must be accounted for if the markedness theory is to meet criterion C is the notion of a « possible alternation ». By « possible alternation » what I have in mind is this: given a phonological rule $A \rightarrow B$, in some context, are there any constraints on the set of features in A and/or B? Let us focus on B. As was noted above (see (4)), phonological rules do not necessarily compound even when their compound would itself be a relatively simple rule. If one considers such « non rules » in the context of the feature hierarchy given by the markedness theory, then one immediately notices that it is in just those cases where rules do not compound that the features specified by the rule, i.e., the features in B, are not linearly ordered. In this respect, conditions on the inputs, i.e., A, to rules differ from conditions on their outputs as specified in the rule, i.e., B. It is not at all unusual to find rules the inputs to which are defined on non linearly ordered sets of features, e.g., [— son, — lab]. The set of possible alternations is characterized in C-II.

Definition:

A *natural class* of segments is the maximal set of segments each member of which is $[\alpha_i, F_i], \dots,$

$[\alpha_j, F_j]$ for any subset of features

$$\mathcal{F} \subseteq \Phi$$

Let $\mu_i = \mathcal{F}_i$ such that for each $k, m, l > k > k+m > j$, F_{i_k} is linearly ordered before $F_{i_{k+m}}$ (i.e., μ_i is a linear sub-path of the feature hierarchy)

C-II Given the phonological rule

$A \rightarrow B$ in some context

(a) A and B are natural classes, and

(b) if B is the non-empty natural class

$\alpha_1 F_1, \dots, \alpha_i F_i$ and $\mathcal{F}_i \subseteq \Phi$, then

$$\mathcal{F}_i = \mu_i$$

From C-II(a) it follows that the inputs and outputs of rules will be null, single segments or sets of n segments, $n > 1$. From C-II(b) it follows that in the last instance the set of segments will be defined on some set of features which are linearly ordered on the feature hierarchy.

Given (C) generative phonological rules of epenthesis compared with other rules that can be answered in terms of markedness. On account of markedness specifications for epenthesis/deletion in languages.

There is a relation between (b) and (c) in a language where the rule cannot be stated in a language where the rule is a natural class in a sense a natural class expressing the idea of a possible alternation postulating a natural class task ever rules, where the child given the demand to hypothesize from the star rules (independent for in terms of the postulate the provides then in the set, and rules up to co

By C-II(b) segments by simultaneous speech those errors in particular the ment which v

Given C-II, or any other approach to alternation within generative phonology, the problem arises as to how one is to evaluate rules of epenthesis and deletion — what is their relative cost compared with other rules of segmental alternation? This is a question to be answered in terms of both the evaluation procedure and markedness. In part the answer will rely on an account of the markedness structure of strings. Leaving aside that component of the account, we will say that the markedness of a segment under deletion or epenthesis is at least in part a function of the number of marked specifications in its expression in a rule. Thus, the markedness of the epenthesis/deletion of a segment *S* will not be constant across languages.

There is a second point to be made with regard to C-II and that relates to (b). By (b) it is not claimed that, e.g., *x* and *n* can't alternate in a language. The claim of (b) is that an alternation of *x* and *n* cannot be stated in terms of a single rule of grammar. So, in Kwakiutl, a language where *x* and *n* do alternate that alternation is captured by two rules: a rule of fronting and a rule of nasalization. C-IIb then is in a sense a component of the evaluation procedure in that by expressing the intrinsic relations of the set of features it entails that certain possible alternations will be quite costly. It should be clear that postulating C-II(b) in no way increases the complexity of the acquisition task even for the case where an alternation requires two or more rules, where one or more is a rule of absolute neutralization. The child given the primary data of the alternation and an analytic demand to hypothesize a rule account for the alternation, tacitly knows from the start that there is a limited and fixed number of possible rules (independent of context); if an alternation cannot be accounted for in terms of one of those rules, the child has no choice but to postulate the simplest set of rules consistent with C-II(b). C-II(b) provides then the independent motivation for all save one of the rules in the set, and the evaluation procedure dictates collapsing of those rules up to consistency with C-II(b).

By C-II(b) there is a rather limited set of possible pairings of segments by phonological rule. In a recent study of normal spontaneous speech errors, Shattuck-Hufnagel and Klatt (1979) consider those errors in which a person substitutes one segment for another, in particular the set of such errors where there is no apparent environment which would lead one to suspect that the error represented

either anticipation or perseveration of some segment in the string. In the corpus they consider there are 1620 nonsyllabic segmental substitutions which are, apparently, neither anticipatory or perseverative. Just considering the surface representation of these errors, strikingly, one notes that 83% of the errors involve pairs of segments which can alternate by a single phonological rule. Given the nature of the data this is a fairly robust correlation. Further analysis indicates that the correlation is probably even greater. The 285 errors (pairs of segments) not accounted for immediately in terms of C-II(b) fall in 31% of the nonrule alternating pairs of (surface) segments in English. That is, the correspondence between impossible rules and nonattested cases is 69%. Thus, in terms of the correspondence of attested cases with possible rules and the correspondence of nonattested cases with impossible rules, the correlations are in the right direction and well above chance. Of the cases which do not correspond to possible phonological rules, one should — « dark » / alternates with labials, phonetically not a surprising alternation but one which is not captured by the feature theory of (9) or (10); 5% of the unexplained errors involve this alternation. For reasons which are totally unclear to me, all stops tend to be substituted for by *f*; 8% of the nonrule errors pairs to be explained involve the substitution of *f* for a stop. Additionally, there are just eight nonrule alternating segment pairs which account for another 23% of the unexplained errors. These three classes of errors, 36% of the cases that are inconsistent with C-II(b), fall on only 5% of the segment pairs where it is predicted by the markedness theory that there can be no single rule alternations. While it is clear that some account is required for these data, it is equally evident that they do not provide very strong counterevidence to the claim that the relation between the intended and produced segment in non-anticipatory/perseverative spontaneous substitutions is governed at least in part by the markedness theory. Finally, it should be noted that of the remaining unaccounted for cases, in 68 (24%) of the cases there is only one attested substitution in the corpus. Given the method of recording spontaneous speech errors, some margin must be allowed for recorder error, and, furthermore, it may well be the case that some of the errors are anticipatory or perseverative (a possibility that Shattuck-Hufnagel and Klatt acknowledge).⁸ That there is such a close correspondence bet-

⁸ The speech error data are discussed in detail in Kean (1979).

ween the coincident mot

It wa
ic) realiza
tion of su
patible w
phonolog
the set of
terpreted
acquisitio
be viewec
Rather, a
segments
Universal
is not dissi
patible wit
in his lang

Havein
have accou
procedure
first the q
this proble
systematic
by rule and
Linking. T
tions do sc
theory AI-I
be stated a:

(15)

between the speech error data and the rule cannot be taken to be a mere coincidence. The speech error data are taken as providing independent motivation for C-II(b).

It was noted earlier that a learning theory for segmental (phonetic) realization would be a « selection » type of learning. The assumption of such a learning theory for segmental realization is also compatible with the markedness theory. Given that all segments are phonologically specified for all features in all representations, then the set of features, expressed by the markedness conventions and interpreted by the Complement Convention, must be constant, that is, acquisition of inventories in their phonological representation cannot be viewed as selection of some subset of the markedness rules. Rather, acquisition involves the selection of that set of phonetic segments which is consistent with experience and C-I mediated by Universal Phonetic Theory. In terms of rule acquisition the situation is not dissimilar. The child must learn which of the set of rules compatible with maintaining C-I, and possible rules by C-II, is exploited in his language.

Having defined the notion « possible alternation by rule » we have accounted for only one of the five « failures » of the evaluation procedure mentioned in 2.1. Let us turn now to the others. Consider first the question of the asymmetry of alternations. In considering this problem Chomsky and Halle (1968) proposed that there was a systematic interaction between the changing of feature specifications by rule and the markedness conventions, this interaction they called Linking. The basic effect of Linking is to let Markedness Conventions do some of the work of phonological rules. In a markedness theory AI-III, using the features of (9), the principle of Linking can be stated as (15).

(15) Given the phonological rule

$A \rightarrow B$ in some context

where: (a) B is a natural class defined on $\mu_{j,k}$

(b) $F_{i_h} \in \mu_{j,k}$, and

(c) there exists no $F_{i_{h+m}} \in \mu_{j,k}$ such

that $F_{i_h} > F_{i_{h+m}}$

for all optional $F_{i_{h+m}}$, $F_{i_h} > F_{i_{h+m}}$, the markedness convention for $[u F_{i_{h+m}}]$ applies to any segment S, S in B, as

if S were specified $[u f_{h+m}]$ just in case S satisfies the environment of the convention.

By (15) the rule of palatalization for $k \rightarrow \check{c}$ is stated:

- (16) $k \rightarrow [-\text{back}]$ in some context

The changes in the specifications of other features is effected through Linking by the markedness conventions, in this cases (17a, d, e).⁹

- (17)
- a. $[u \text{ cor}] \rightarrow [+ \text{ cor}] \quad / \quad \begin{bmatrix} -\text{syl} \\ -\text{back} \\ +\text{lab} \end{bmatrix}$
- b. $[u \text{ spr}] \rightarrow [+ \text{ spr}] \quad / \quad \begin{bmatrix} -\text{syl} \\ -\text{cor} \\ -\text{back} \\ +\text{lab} \end{bmatrix}$
- c. $[u \text{ hi}] \rightarrow [-\text{ hi}] \quad / \quad \begin{bmatrix} -\alpha\text{CONS} \\ \alpha\text{ant} \\ \beta\text{cor} \\ -\beta\text{lab} \\ -\beta\text{spr} \end{bmatrix}$
- d. $[u \text{ DR}] \rightarrow [+ \text{ DR}] \quad / \quad \begin{bmatrix} -\text{son} \\ +\text{cont} \\ +\text{cor} \\ \alpha\text{hi} \\ -\alpha\text{lat} \\ -\text{back} \end{bmatrix}$
- e. $[u \text{ stri}] \rightarrow [+ \text{ stri}] \quad / \quad \begin{bmatrix} -\text{son} \\ \alpha\text{cont} \\ -\alpha\text{DR} \\ -\text{back} \\ -\text{SG} \\ -\text{lat} \end{bmatrix}$

⁹ van Lessen Kloeke (1981) provides considerably more elegant markedness conventions than those in (17) which are taken from Kean (1975).

By co
would

(18

Th
tions, (t
should
the pro
now a
blems (t
It i
stops, (t
C-II. F
lenition
tively w
tions an
problem
are voic
evaluati
tion; un
is, howe
spirantiz
rules are
count fo
tion with
Karak. T
in most t
consider
phonetic
place of t

¹⁰ It v
becomes un
tion of the c
phonetic rea

By contrast, to state the alternation $\check{c} \rightarrow k$, under (15), the rule (18) would be required

$$(18) \quad \check{c} \rightarrow \left[\begin{array}{l} + \text{back} \\ -\text{cor} \\ -\text{DR} \\ -\text{stri} \end{array} \right] \quad \text{in some context}$$

Therefore, given the hierarchy of features/markedness conventions, (15), and the evaluation procedure, it is predicted that $k \rightarrow \check{c}$ should not alternate symmetrically. Thus, in addition to « solving » the problem noted in (1), Linking « solves » the problem (2), $k \rightarrow \check{c}$ is now a rule changing one feature.¹⁰ Linking also addresses the problems (3) and (5).

It might have been noted in passing that the lenition of voiceless stops, (5) above, cannot apparently be stated as a single rule under C-II. For example, under standard accounts the β and γ derived by lenition are fricatives; to derive these segments from p and k , respectively would require two rules — one to alter the voicing specifications and one to alter the continuance specifications. There are two problems with the assumption that the outputs of such lenition rules are voiced non strident fricatives. First, there is a failure of the evaluation in the sense that lenition is a very costly sort of alternation; under anyone's theory far more costly than spirantization. This is, however, inconsistent with the observation that lenition, like spirantization, is quite common. Secondly, if the outputs of lenition rules are fricatives and not sonorants, there is no natural way to account for the fact that such derived segments pattern under alternation with sonorant continuants and not fricatives, as in, for example, Karok. Third, there is the curious fact that nondental t 's lenite, as t 's in most cases (e.g., the Numic languages) and t 's, otherwise. These considerations, taken together with the fact that for no principled phonetic reason there is not in most accounts a full complement by place of nonlateral liquids strongly suggests that the segments deriv-

¹⁰ It will be noted that given these markedness conventions fronting of k to \check{k} , becomes unacceptably costly. van Lessen Kloeke overcomes this through reformulation of the conventions. My current approach is to view this alternation in terms of the phonetic realization features rather than phonological features.

es the en-

state i:

ed through
, d, e)?

1
ick
b

1
or
ack
ab

cons
ant
cor
lab
spr

son
cont
cor
ahi
lat
back

-son
cor.t
-αDF
-back
-SG
-lat

giant markedness

ed by lenition are in fact sonorants.¹¹

Given this reinterpretation of lenition we can state the rule of lenition as (19).

$$(19) \quad \left[\begin{array}{c} -\text{cont} \\ +\text{StVC} \end{array} \right] \rightarrow [+ \text{son}] \quad \text{in some context}$$

The markedness conventions for *StVC* and *cont* will apply to derive the liquids.

$$(20) \quad \begin{array}{l} [\text{u cont}] \rightarrow [+ \text{cont}] \quad / \quad \left[\begin{array}{c} + \text{son} \\ - \text{nas} \end{array} \right] \\ [\text{u StVC}] \rightarrow [- \text{StVC}] \quad / \quad \left[\begin{array}{c} + \text{son} \end{array} \right] \end{array}$$

This analysis, it should be obvious, shows that the apparent problems (3) and (5) are not problematical at all. It might also be noticed that this analysis provides yet another argument against the dictum « same sound same symbol » — though it is doubtful that more such are needed given Halle (1959).

While (15) does solve problems, the fact remains that its effectiveness is contingent on the acceptance of the capricious notion of an optional feature. Under the hypothesized limited and abstract feature system, no such notion is required, Linking is stated as C-III.

C-III Linking

Given the rule

A → B in some context

where: (a) B is the natural class defined on $\mu_{j,k}$

(b) $F_h \in \mu_{j,k}$, and

(c) there exists no $F_{h+m} \in \mu_{j,k}$
such that $F_h > F_{h+m}$

¹¹ In Kean (1975) it is suggested that the segments derived by lenition are glides. In my work in developing a new feature system and in van Lessen Kloetke's work using an SPE type feature system it has become clear that the segments derived by lenition cannot be glides and must be liquids.

C-III
of « o

3. Son

3.1 Th

It was t
agreem

B.

While it
will also
equal, t
and C, i
pected t

A.

C.

That is,
satisfyin
rently sa
expressic
redundan
grounds
least in p

From
develop a
foundatic

for all $F_{h+m}, F_h > F_{h+m}$, the markedness convention for F_{h+m} applies to any segment S, S in B, as if S were specified [u F_{h+m}] just in case S satisfies the environment of the convention.

C-III differs from (15) only in that it makes no appeal to the notion of « optional » features.

3. Some Summary Consideration

3.1 The Interdependence of A, B, and C

It was taken as an assumption at the outset that there would be general agreement that any markedness theory would satisfy criterion B.

- B. A theory of markedness will assign probabilities of occurrence and variation of substantive elements within and across grammars

While it is not *a priori* obvious to assume that any theory satisfying B will also satisfy A and C, it was argued that all other things being equal, given a theory meeting B and the need for UG to include A and C, in the absence of any evidence to the contrary it was to be expected that a markedness theory meeting B would also meet A and C.

- A. A theory of markedness is a formal theory of substantive universals
 C. A theory of markedness includes non parameterized constraints which provide the substantive definition of possible grammar for human language

That is, while it is not logically necessary that a markedness theory satisfying B should also satisfy A and C, if A and C are not concurrently satisfiable with B, then there will be a redundancy in UG in the expression of substantive elements. Without motivation for such redundancy, one would then be forced on purely methodological grounds to assume that A, B, and C characterized the properties (at least in part) of any theory of markedness.

From this rather general point it follows that any attempt to develop a markedness theory over any general domain will have at its foundation some explicit attempt to meet A. To make any proposal

ie rule of

text

/ to derive

son
nas

son

erent prob-
be noticed
the dictum
t more such

at its effec-
is notion of
nd abstract
ed as C-III.

$\mu_{i,k}$

emission are glides.
seke's work using
erived by lenition

for a markedness theory which satisfies B but not A would, of course, have no bearing on the empirical content of the hypothesis that a markedness theory must satisfy A and B. By the same token, a markedness theory satisfying B and not C would also have no bearing on the hypothesis that a markedness theory must and will meet C as well as B. It will always be possible to stipulate the set of substantive universals for any domain, and then place arbitrary constraints on the arraying of those elements and thereby provide a substantive definition of possible human languages; therefore any such proposals cannot in themselves vitiate A-C.

The theory of markedness characterized by A-II, II, III, B-I, II, and C-I, II, III is a theory which satisfies the criteria A, B, C. This theory derives its motivation from the fact that it provides a principled account for the intrasegmental phenomena which any markedness theory must account for if it is to meet B and C. It provides *prima facie* evidence for the hypothesis that a theory of markedness will meet A, B, C in that it is a markedness theory meeting B which at the same time also meets A and C. It has, I think, just the sort of properties that one would expect any such theory to have. To take the three obvious classes of examples.

(i) If the assumption that a markedness theory must meet both B and C were false, then it would be expected that in any attempt to construct a theory meeting B and C that there would be no interaction between those two aspects of the theory, that is, that all evaluative considerations bearing on variation would fall under B. However, as should be evident in the case of the markedness theory outlined here, that is not the case. The data presented in (1-5) are all problems of evaluation, problems of the relative likelihood of various intrasegmental alternations. If B and C were independent in UG there would be no reason to expect that any of those data be accounted for through an interaction of the evaluation procedure and the substantive characterization of possible grammar. However, C-II, III are conditions on grammars, not languages, and they do, taken together with the evaluation procedure provides an account for a class of rule variation. In discouraging segmental complexity in markedness specifications, B-I, II contribute a probabilistic evaluative constraint across the derivations in a grammar. Thus, the characterization of variation involves a conspiracy between the B and C components of the system.

(ii)
ponents
constrain
demand
participate
constraints
normal theo
rather su
massive
vious wh
tories of
striking
stated in

(iii)
tem meet
that any c
reconsider
implausib
out havin
conventio
the explan
rests on t
such a sys
expressed.
substantiv
must be m
then a cha
pected if a
In adoptin
have been
features (c
should ari
have been

Simila
tive domain
of such in
richness of

(ii) A second case of interaction between the B and C components involves B-I, II, and C-I. There is a tension between these constraints; while complexity is costly under B it is in some cases demanded by C-I. C-I itself has a property we would surely anticipate in a markedness theory satisfying criteria C and A. If constraints on possible segmental systems were not based on some formal theory of the set of features, then it would not be expected that rather small differences in segmental systems would have rather massive superficial consequences. For example, it is not at all obvious why there should be such differences in the consonantal inventories of two vowel vs. three vowel systems. That one does find such striking differences follows immediately from the fact that C-I is stated in terms of the array given by A-I, II.

(iii) The tensions and interactions between components of a system meeting A, B and C would be expected to have the consequence that any change in the characterization of one component would force reconsideration of every other component. For example, it would be implausible to assume that one could change either A-I or A-II without having to radically change the entire theory. As the markedness conventions expressed under A-I express a substantive domain, and the explanatory force of the B and C components of such a system rests on those specific expressions, it is also to be expected that in such a system questions will be raised as to the set of elements to be expressed. That is, phonological features not only need be motivated substantively and classificatorily (in the traditional sense) they also must be motivated in the context of the markedness theory. There is then a change in what it means to motivate a feature as would be expected if a theory of segmental structure were to meet A, B, and C. In adopting the markedness theory outlined in section 2, questions have been raised about the motivation and interpretation of various features (cf., van Lessen Kloeke, 1979, 1981); that such questions should arise is to be expected given that the criteria of motivation have been altered.

Similar situations would be expected to obtain in other substantive domains. It should be evident of course that the effective force of such interactions will be contingent in some measure on the richness of the substantive vocabulary of a domain.

3.2 Markedness and Learning

The evaluation procedure (simplicity metric) distinguishes among descriptively equivalent descriptively adequate grammars. Given a formal theory of rules, the evaluation procedure distinguishes among possible descriptively adequate grammars for some language L, by the simple expedient of counting the substantive elements which are expressed in the rules of those grammars. The empirical content of the evaluation procedure rests on the formal theory that is postulated; different formal theories will make different demands on the expression of substantive elements in rules. Thus, the evaluation procedure, while operating on a count of substantive elements, can be conceived of distinguishing descriptively equivalent grammars in virtue of the theory of formal universals. It takes no account of the substantive structure of grammars *per se*. Its function in an account of language acquisition is to determine which of a set of competing potential grammatical hypotheses it is that a child will entertain.

All grammars for possible human languages must have a substantive vocabulary. A child cannot be conceived of as simply acquiring the rule system of the grammar of his language and not the substantive vocabulary over which that rule system operates. A child cannot acquire a rule system in the absence of substantive acquisition. Given that there is substantive variation among languages, the child may entertain a variety of (potential) substantive vocabularies for his language in the course of acquisition. If there were no means for evaluation the possible substantive hypotheses which a language learner might entertain, then there would be no reason why it could not be the case that, e.g., in the ludicrous extreme, an English speaker might postulate a systematic phonemic inventory including all the segments of English plus λ . The evaluation procedure cannot exclude this possibility because it operates over rule systems. Given that for any language there can be segments which do not function as factors in rules, it cannot be the case that the segmental inventory for any language is determined solely on the basis of what is required for the rule system to function. Given this one might be led to assume either that the child builds up a segmental inventory incrementally adding one or two segments at a time and modifying previous segmental hypotheses, or that he assumes the set of all possible segments and then pares that set down to the system of his language on the basis of his experience, tossing out those segments for which

he has allowed for human approach. Neither substantive language domain

Just (tions) th acquisition and the cedure (is concu entertain vocabulari: child's j evaluati strain a language straightf cessive s B'.

By B' it simplest cedure, s elements be equal theory G₁ tion proc necessary a child w

Ther markedne

he has no evidence. Under the former approach the child will be allowed to entertain « grammars » which are not possible grammars for human languages. The most coherent interpretation of the latter approach is under an assumption of instantaneous acquisition. Neither approach is constrained by any notion of there being a substantive definition of segmentally possible grammar for a human language. Analogous arguments can be made for other substantive domains.

Just as it is plausible to assume (under the appropriate idealizations) that a child entertains a sequence of grammars in the course of acquisition, his successive hypotheses constrained by his experience and the theory of possible grammars including the evaluation procedure (simplicity metric), it is also plausible to assume that the child is concurrently acquiring the substantive vocabulary of his language, entertaining successive hypotheses as to the appropriate substantive vocabulary of his language. Just as it is necessary to constrain the child's grammatical hypotheses in terms of the operation of the evaluation procedure over its domain, so too is it necessary to constrain a child's hypotheses of the substantive vocabulary of his language. The theory of markedness presented here offers a straightforward means for providing an account for a child's successive substantive grammatical hypotheses:

- B'. Given a sequence of grammars, $G_0, \dots, G_i, \dots, G_L$ which a child entertains in the course of acquisition of the grammar G_L for the language L , the substantive vocabulary for G_i is the least marked possible substantive vocabulary (consistent with his experience).

By B' it is simply claimed that just as the child hypothesizes the simplest rule system available on the basis of the evaluation procedure, so too he hypothesizes the least marked substantive set of elements for that grammar. Under the evaluation procedure G_{i+1} may be equally simple or more complex than G_i ; under the markedness theory G_{i+1} may be equally or more marked than G_i . Both the evaluation procedure (simplicity metric) and the markedness theory provide necessary constraints on the successive grammatical hypotheses that a child will entertain given his experience.

There will be an interaction and a tension between the markedness theory and the evaluation procedure (simplicity metric)

in their evaluative function in the determination (selection) of the most highly valued (least marked and simplest) descriptively adequate grammar as is illustrated by the following examples:

(i) Consistent with his experience, the child may entertain the minimally marked substantive vocabulary S for his grammar or the equally marked but distinct substantive vocabulary S' ; given S he must hypothesize the set of rules R , and given S' , R' , where R is less complex, i.e., simpler, than R' by the simplicity metric. Under such circumstances, given B' and the simplicity metric, the child hypothesizes S and R .

(ii) Consistent with his experience, the child may hypothesize the set of rules R or the set of rules R' , where R and R' are equally valued by the simplicity metric. However, given R the child must hypothesize the substantive vocabulary S , and given R' he must hypothesize the substantive vocabulary S' , where S and S' are not equally marked, S being less marked than S' . Under such circumstances, given B' and the simplicity metric, the child hypothesizes S and R .

As with the theory which postulates just the simplicity metric, there will be cases under the approach which adopts both B' and the simplicity metric which will not be decidable. For example, it is possible that a child might arrive at some point where consistent with his experience there is more than one grammar available to him where those grammars are of equal simplicity and of equal markedness. Holding simplicity constant, given the potential variation in markedness, it should be evident that with the adoption of B' that the set of undecidable cases which may arise has been significantly reduced, that is, B' further constrains the theory of the course of grammatical acquisition by reducing the potential class of cases which are undecidable on the basis of the evaluation procedure alone. There is a second class of cases where the adoption of B' raises a question of decision which does not arise when one adopts the simplicity metric alone. It is conceivable that a child might arrive at some point where consistent with his experience he might postulate either G_i or G'_i ; the rule system of G_i is simpler than the rule system of G'_i , but the substantive vocabulary of G_i is more highly marked than that of G'_i . It is unclear what one would want to say about such cases.

3.3 Mar

A very c
learning
They ma

In their c
ed than C
 G_i of so
grammar
is consi
evaluativ
members
not, hov
which is
proposed

For
successiv
grammar
that G_0
substanti
acquisitic
not be in
substanti
ever, that
tem does
substanti
the claim
consisten
not differ
distinguis

(i) T
same grat

3.3 *Markedness and Attainability*

A very different view of markedness, in general, and with respect to learning, in particular, is adopted by Rouveret and Vergnaud (1980). They make the following proposal.

The complexity of the grammar G_L in terms of markedness can be equated with the number of changes which have been made in the set of rules G_0 [the initial grammar] to get the descriptively adequate grammar for L [a language], G_L , i.e., with the length of the sequence G_0, \dots, G_L (p. 195).

In their discussion they note that, « by this definition G_i is less marked than G_{i+1} , even if the two grammars only differ by the addition to G_i of some mechanism drawn from Core grammar », where core grammar is regarded « as the stock of hypotheses of null cost ». There is considerable appeal to their suggestion that there is some evaluative distinction between a sequence of grammars with m members and a sequence with n members, $m \neq n$. Their proposal is not, however, a learning theoretic interpretation of markedness which is compatible as stated with the theory of markedness theories proposed here.

For their hypothesis to make sense it must be assumed that each successive grammar is a well-formed grammar, that is, a possible grammar of a human language. This being the case, it must be the case that G_0 is under their view substantively well-formed, i.e., both substantively and formally a possible grammar. Therefore, while the acquisition of the rule system of the grammar of some language need not be instantaneous under their view, there must be instantaneous substantive acquisition up to well-formedness. It should be clear, however, that instantaneous acquisition of a well-formed substantive system does not entail instantaneous acquisition of the appropriate substantive system for the language the child is learning; it only makes the claim that the child hypothesizes a potential substantive system consistent with his experience. In this the two views of acquisition are not different. Consider now the following four possibilities which do distinguish the two approaches.

(i) There are two children, α and β , each of whom arrives at the same grammar (i.e., substantive and formal equivalence). Due to dif-

fering environments, one child hypothesizes G_{O_α} while the other hypothesizes G_{O_β} , $G_{O_\alpha} \neq G_{O_\beta}$. Let us further assume that each child goes through an equal number of stages in the course of acquisition, that is, $G_{L_\alpha} = G_{n_\alpha}$ and $G_{L_\beta} = G_{n_\beta}$, n_α and n_β successive positive integers assigned to each successive grammar, G, G_0, \dots, G_n , $n_\alpha = n_\beta$, then under the Rouveret-Vergnaud hypothesis G_{L_α} is as marked as G_{L_β} . Their interpretation leaves open no possibility for saying anything about the relative markedness of G_{O_α} versus G_{O_β} ; in this it is distinct from the view of markedness given in A, B, and C since under such theories G_{O_α} and G_{O_β} will be more or less marked with respect to each other.

(ii) Consider the children, α and β , who arrive through the course of acquisition at distinct grammars. As a consequence of experience one child postulates G_{O_α} and the other G_{O_β} ; in this case, however, $G_{O_\alpha} = G_{O_\beta}$. Through the course of acquisition the children go through the same number of stages (grammars), i.e., $G_{L_\alpha} = G_{n_\alpha}$ and $G_{L_\beta} = G_{n_\beta}$, $n_\alpha = n_\beta$, but here $G_{L_\alpha} \neq G_{L_\beta}$. Again under the Rouveret-Vergnaud account G_{L_α} is as marked as G_{L_β} . But under the hypothesis as to the structure of markedness theories given here, while G_{O_α} is as marked as G_{O_β} , it is not necessarily the case that G_{L_α} is as marked as G_{L_β} (though they may be equally marked).

It might be suggested that (i) and (ii) are not well-formed examples illustrating the distinctness of the Rouveret and Vergnaud position from that adopted here on the grounds that it should be assumed that for all children there is but one initial grammatical state G_0 . If, however, we assume that the initial grammar (and any subsequent grammar) a child entertains is a possible grammar of a human language then for any child, G_0 , his initial grammar (not his initial state), will be a function of his experience *modulo* his tacit knowledge of possible grammars, i.e., UG; this being so there could not be a single G_0 for all children. It makes little sense to conceive of the child as passing through a sequence of grammars in the course of acquisition and not assume that each of those grammars is a possible grammar for a human language. UG itself is not a possible grammar for a human language, rather it sets the parameters of what grammars can be attained. The Rouveret and Vergnaud proposal would be uninterpretable were one not to assume that each grammar that a child hypothesizes is a possible grammar. As was argued in Section 1,

as all g
vocabul
mar wi
degree

(iii)

gramme
due to c
 G_0 whil
rive at t
number
case the
that G_L
Howeve
marked

(iv)

$\neq G_{L_\beta}$; f
Rouvere
markedr
of the s

It sl
markedr
are incor
acquisiti
acquisiti
evaluativ
hypothes
the same
states wh
in terms
the numl
the cours
of the R
views ad
the exten
terminolo
that the t
the issue
« marke

as all grammars of human languages must contain a substantive vocabulary, if G_0 for any child is a possible grammar then that grammar will have a substantive vocabulary and, consequently, some degree of markedness under the position adopted here.

(iii) Again, take two children, in this case they arrive at the same grammars $G_{L_a} = G_{L_b}$. Starting from initial equivalence, $G_{O_a} = G_{O_b}$, due to differing experience one child arrives at the intermediate state G_{I_a} while the other arrives at G_{I_b} , $G_{I_a} \neq G_{I_b}$. While they ultimately arrive at the same final state, the children progress through a distinct number of stages, i.e., $G_{L_a} = G_{n_a}$, $G_{L_b} = G_{n_b}$, but $n_a \neq n_b$. In this case the Rouveret and Vergnaud hypothesis would make the claim that G_{L_a} was more or less (but not equally) marked than G_{L_b} . However, under the markedness theory given here G_{L_a} would be as marked as G_{L_b} .

(iv) Two children acquire distinct grammars, G_{L_a} and G_{L_b} , $G_{L_a} \neq G_{L_b}$; furthermore, $G_{L_a} = G_{n_a}$, $G_{L_b} = G_{n_b}$, but $n_a \neq n_b$. Following Rouveret and Vergnaud G_{L_a} and G_{L_b} are necessarily not of equal markedness, but under the proposal made here G_{L_a} and G_{L_b} may be of the same markedness (though they aren't necessarily).

It should be evident from these cases that the two approaches to *markedness*, that of Rouveret and Vergnaud and that presented here, are inconsistent. If one accepts the position that there is substantive acquisition as well as the acquisition of a rule system involved in the acquisition of a grammar, and furthermore that there must be some evaluative means of distinguishing between competing substantive hypotheses then something along the lines of B' will be required. By the same token, if one wants to distinguish the final grammatical states which children (even speakers of the same language) arrive at in terms of their attainability, where attainability is taken to mean the number of successive grammars which the child will entertain in the course of acquisition of some G_L , then something along the lines of the Rouveret and Vergnaud hypothesis will be required. The two views address distinct issues in the theory of acquisition; therefore, to the extent that they are inconsistent this would seem to be a matter of terminology rather than content. For no other reason than the fact that the theory of markedness put forward here addresses (some of) the issues which have by tradition been treated under the rubric « markedness », we will restrict the use of the term *markedness* to

theories of the type discussed here; the term *attainability* will be used with respect to approaches such as that of Rouveret and Vergnaud. It should be evident that drawing such terminological distinctions is not a trivial matter; it is a matter making precise the circumscription of domains, no more trivial than determining the technical interpretation of grammar within the theory of transformational generative grammar, or, at another level, fixing explicit formalisms to make particular claims precise. Given this terminological distinction, the Rouveret and Vergnaud proposal is restated as:

Attainability

The complexity of a grammar G_L in terms of attainability can be equated with the number of changes which have been made in the set of rules and substantive vocabulary G_0 to get the descriptively adequate grammar for L , G_L , i.e., the length of the sequence G_0, \dots, G_L .

To paraphrase Rouveret and Vergnaud (p. 9), by this definition G_i is more attainable than G_{i+1} , even if the two grammars only differ by the addition to G_i of some mechanism drawn from core grammar, where we take « core grammar » to be the set of minimally (least) marked and simplest possible grammars. As Rouveret and Vergnaud note G_i and G_{i+1} may be equally valued by the simplicity metric; we have added here that G_i and G_{i+1} may be equally valued in terms of markedness as well. Neither the simplicity metric nor the markedness theory can evaluate a sequence of grammars; if the claim of Rouveret and Vergnaud that the theory of grammar must include a mechanism for evaluating a sequence of grammars is accepted, Attainability must be adopted. For any $G_i, G_j, G_i \neq G_j$, if G_i falls within the set of core grammars and G_j does not, then G_i will be less complex in attainability (as well as simplicity and markedness) than G_j . For any two grammars $G_i, G_j, G_i \neq G_j$, the attainability of both G_i and G_j will be a function of the course of acquisition which leads the child to arrive at G_i or G_j . In terms of attainability, the optimal course of acquisition would, of course, be instantaneous acquisition; for all other cases the attainability of any grammar G_i will be a function of the sequence of the child's experience, experience taken as a pair consisting of a current grammar G_c and some new data, D , not consistent with G_c which forces the child to hypothesize a new grammar, G_d .

3.4 *Conclusion*

The theory of evaluation theory, and a that the neces here. Further grammars attain rise to them h a notion of att ting attainabi substantive by the theory of independent c is the motivat sic content of mars, and in substantive el theory meetin « markedness *per se* (that is whether there which is called

For any phonological that domain, and, addition: grammars of what can be a subset of the possible set o language; that is making the pressible in te It has been ar

In the ab for two distin terms of G_i co if R_i and R_j c assuming G_i :

3.4 Conclusion

The theory of grammar as it is conceived of here contains then a set of evaluation procedures: the simplicity metric, the markedness theory, and a measure of attainability of grammars. It is assumed that the necessity for the simplicity metric need not be argued for here. Furthermore, the suggestion that there is a need to evaluate grammars attained in terms of the sequence of grammars which gives rise to them has been accepted without argument; on this assumption a notion of attainability has been adopted. The motivation for adopting attainability rests here in distinguishing terminological from substantive hypotheses as to role and function of markedness within the theory of grammar. The justification for attainability is entirely independent of any claims made here. What has been of interest here is the motivation of the need for a theory which expresses the intrinsic content of the substantive elements available for possible grammars, and in doing so evaluates the set(s) of grammatically possible substantive elements. The theory of substantive universals (i.e., the theory meeting A, B, and C) within in UG has been designated « markedness theory ». What is of consequence is not the terminology *per se* (that is historical and of no grammatical interest) but rather whether there must be a theory of substantive universals of the type which is called here a markedness theory.

For any systematic set of substantive elements (e.g., the phonological distinctive features) the theory of markedness expresses that domain, i.e., establishes a domain in which there is acquisition, and, additionally, in providing a substantive definition for possible grammars of human languages, the theory of markedness specifies what can be acquired. In the absence of a theory of markedness, any subset of the set of substantive elements available for grammars is a possible set of substantive elements for the grammar of a human language; that is to say, in the absence of a theory of markedness one is making the empirical hypothesis that substantively anything expressible in terms of the set of substantive elements can be acquired. It has been argued here that such an hypothesis is empirically false.

In the absence of a theory of markedness, given two grammars for two distinct languages, G_i and G_j , if G_i differs from G_j simply in terms of G_i containing the rule R_i and G_j containing R_j , i.e., $R_i \neq R_j$, if R_i and R_j contain the same number of substantive factors, then, assuming G_i and G_j to be of equal attainability, there will be no

evaluative distinction between G_i and G_j . It is an empirical question whether or not it is necessary for such an evaluative distinction to be made; it has been concluded here that such an evaluative distinction must be made. The theory of markedness being the theory of substantive universals provides the means for making such distinctions.

Under the theory of markedness that has been proposed, core grammar consists of the set of maximally simple and minimally marked possible grammars. Given this view, the set of possible departures from core grammar which may be realized in particular grammars is increased to include not only departures from maximally simple rule systems, but also departures from minimally marked substantive vocabularies. If one does not assume that the delimitation of the set of possible grammars in UG includes a claim as to substantively possible grammars, then there is no justification for this interpretation of core grammar. However, if the arguments outlined here are accepted, then the interpretation of core grammar adopted here must be accepted. Just as there can be no possible grammar (a core grammar or other) which has no rules at all (i.e., is infinitely simple), so too it is claimed that there can be no possible grammar which is substantively unmarked.

APPENDIX I

The Features

seg	segment	nas	nasal
syl	syllabic	ant	anterior
cons	consonantal	back	back
son	sonorant	low	low
StVC	stiff focal cords	lab	labial
SIVC	slack vocal cords	cor	coronal
CG	constricted glottis	spread	spread
tense	tense	hi	high

long

SG

cont

With
Sound Pa

1. sʒ

sʒ

fr

cc

2. st

st

St

3. sʒ

se

sʒ

th

is:

sic

Ri

4. de

[+

th:

cro

str

APPENDIX II

 F_i = a phc

 f_i = a pho

 Φ = { F_i, \dots }

 Φ' = { f_i, \dots }

Possible phc

Pos (F_i)

long	long	lateral	lateral
SG	spread glottis	DR	delayed release
cont	continuant	stri	strident

With the exceptions noted below these features are as in the *Sound Pattern of English*.

1. *syllabic*: This feature « replaces » vocalic. All vowels and syllabified sonorant constants are [+ syllabic]; stops, fricatives, affricates, glides, and nonsyllabified sonorant consonants are [- syllabic].
2. *stiff vocal cords, slack vocal cords, constricted glottis, spread glottis*: These features are proposed by Halle and Stevens (1972).
3. *spread*: This feature is introduced in Kean (1975). Labial segments are either [+ spread], i.e., unrounded, or [- spread], i.e., rounded. van Lessen Kloeke (1981) argues that the feature *round* rather than *spread* should be used. At issue between the two features besides markedness considerations is the question of the characterization of the Russian [i] and the Japanese /u/.
4. *delayed release*: In SPE both fricatives and affricates are [+ DR]; following Kean (1975) it is taken here to be the case that only affricates are [+ DR]. In that system the only cross-classification of fricatives and affricatives is between strident members of those sets.

APPENDIX II

F_i = a phonological feature

f_i = a phonetic feature

$\Phi = \{ F_1, \dots, F_n \}$ (the set of phonological features)

$\Phi' = \{ f_1, \dots, f_n \}$ (the set of phonetic features)

Possible phonetic realizations of F_i :

$\text{Pos}(F_i) \subseteq \Phi'$

R_{ij} = rule for the phonetic realization of F_i in context C_{ij}

$$R_{ij} = \pm F_i - \{s_{ij}(f_1), \dots, s_{ij}(f_{n_i}) \mid C_{ij}\}$$

$s_{ijk} = \pm, \neq, +$ or $-$ (the signs or specifications)

Sp = specified phonetic feature(s)

$$\text{Sp}(F_i) \subseteq \text{Pos}(F_i)$$

$$\text{Sp}(F_i) \neq \phi \quad (\text{i.e., } n_i \geq 1)$$

For any R_{ij} and R_{ik} , if $C_{ij} = C_{ik}$, then $R_{ij} = R_{ik}$ (i.e., there is one rule per context)

For L , a language:

I. Disjointness

$$\text{For } i \neq j, \text{Sp}(F_i) \cap \text{Sp}(F_j) = \phi$$

By (I) a phonetic feature f_i may be realized by at most one phonological feature F_i

II. Completeness

$$\varnothing$$

$$\bigcup_{i=1} \text{Sp}(F_i) = \Phi'$$

$$i=1$$

Every phonetic feature must be realized.

From I and II it follows that any f_i is realized by exactly one F_i . I and II mean the sets of specified phonetic features form a partition of the phonetic feature.

III. Distinctiveness

dfn. f_{ij} is *distinctive* in R_{ik} iff s_{ijk} is \pm or \neq

For every i , there exists a j, k s.t.

f_{ij} is distinctive on R_{ik}

By III, every phonological feature has a distinctive realization.

Φ is a partially ordered set, that is, there is a hierarchy of features.

$\mathcal{F}_i = F_1, \dots, F_m$ is a linear set of features if F_1, \dots, F_m is a linear subpath of the feature hierarchy

For any C_{ij} there is an \mathcal{F}_{ij} , linear, s.t. C_{ij} = the conjunction of \mathcal{F}_{ij} signed with $+$ or $-$.

C_{ij} may be the empty set

IV. Nestedness

$$\forall_{i,j,k}, \mathcal{F}_{ij} \subseteq \mathcal{F}_{ik} \text{ or } \mathcal{F}_{ik} \subseteq \mathcal{F}_{ij}$$

The set of all contexts is linear

V

VI.

VII.

VIII.

References

- Cairns, C.
cy F
Chomsky,
brid
Chomsky,
Chomsky,
Chomsky,
and

- (i) for any i , there is a maximal \mathcal{F}_i
 (ii) for any i , $\cup_j \mathcal{F}_i$ is linear (and, in fact, = the maximal \mathcal{F}_i).

V. Order

If $\mathcal{F}_i \subsetneq \mathcal{F}_k$, then R_k precedes R_i ($R_k < R_i$)
 V implies that if $\mathcal{F}_i = \mathcal{F}_k$, then R_i and R_k are unordered.

VI. Exhaustiveness

For every F_i , there is an R_j s.t. $C_{ij} = \phi$
 From VI it follows that F_i is realized in every segment of L_i

VII. D-Consistency

If $R_i \leq R_k$, and if f_{ik} is distinctive in R_k then $s_{ik} = s_{ijk}$.

The effect of VII is that if f_{ik} is specified \pm in R_j then it must also be specified \pm in R_k .

From VII it follows that if f_{ik} is not distinctive in R_j , then f_{ik} is not distinctive in R_k (the contrapositive of VII)

VII may be too strong for the case of $R_j = R_k$; that is, VII may hold only for $R_j > R_k$, and a separate condition, VII', holds for $R_j = R_k$.

VII': If $R_j = R_k$ and if f_{ik} is distinctive in R_j and R_k , then $s_{ijk} = s_{ik}$.

VIII. N-Consistency

If f_{ik} is not distinctive in R_j and R_k , then $s_{ijk} = s_{ik}$.

From VIII, it follows that if f_{ik} is specified, e.g., + in R_j then it cannot be specified — in R_k .

References

- Cairns, C.E. (1969) "Markedness, Neutralization, and Universal Redundancy Rules," *Language* 45, 863-885.
 Chomsky, N. (1965) *Aspects of the Theory of Syntax*, MIT Press, Cambridge Mass.
 Chomsky, N. (1980) "On Binding," *Linguistic Inquiry* 11, 1-46.
 Chomsky, N. (1981) "Markedness and Core Grammar," this volume.
 Chomsky, N. and M. Halle (1968) *The Sound Pattern of English*, Harper and Row, New York.

s one rule

t most one

exactly one
 tic features

alization.
 hierarchy of

if $F_i, \dots,$

conjunction of

- Chomsky, N. and H. Lasnik (1977) "Filters and Control," *Linguistic Inquiry* 8, 425-504.
- Dresher, E. (1979) "Rule Markedness and the Acquisition of Abstract Phonologies," Abstract, *GLOW Newsletter*, 2.
- Halle, M. (1959) *The Sound Pattern of Russian*. Mouton, The Hague.
- Halle, M. and K.N. Stevens (1972) "A Note on Laryngeal Features," *Quarterly Progress Report of the Research Laboratory in Electronics*, MIT, # 101, 198-213.
- Hubel, D.H. and T.N. Wiesel (1970) "The period of Susceptibility to Physiological Effects of Unilateral Eye Closure in Kittens," *Journal of Physiology* 160, 106-154.
- Jakobson, R., G. Fant and M. Halle (1952) *Preliminaries to Speech Analysis*. MIT Press, Cambridge.
- Kean, M-L (1975) *The Theory of Markedness in Generative Grammar*. Unpublished doctoral dissertation, MIT.
- Kean, M-L (1977) "'Natural Processes' and 'Learned Rules' in Markedness Theory," *NELS VII*, 135-146.
- Kean, M-L (1978) "Some Features of Features," *GLOW* 3, University of Amsterdam.
- Kean, M-L (1979) "The Crucial Role of Distinctive Features and Markedness in Speech Production," unpublished ms., University of California, Irvine.
- Lasnik, H. and R. Freidin (1981) "Core Grammar, Case Theory, and Markedness", this volume.
- Rouveret, A. and J-R Vergnaud (1980) "Specifying Reference to Subject: French Causatives and Conditions on Reference," *Linguistic Inquiry* 11, 97-202.
- Sapir, E. (1925) "Sound Patterns in Languages," *Language* 1, 37-51.
- Shattuck-Hufnagel and D. Klatt (1979) "The Limited Use of Distinctive Features and Markedness in Speech Production: Evidence from Speech Error Data," *Journal of Verbal Learning and Verbal Behavior* 18, 41-56.
- van Lessen Kloeke, W.U.S. (1979) "The Markedness of Laryngeals," in W. Zonneveld (ed.) *Linguistics in the Netherlands 1977-1979*, Foris, Dordrecht.
- van Lessen Kloeke, W.U.S. (1981) "How Strident is the Raspberry? Likely, Unlikely, and Impossible Feature Configurations in Phonology," this volume.

M.-L. Kean
 University of California
 Irvine, California 92717

2318 001