Faculty Research and Creative Activity   Communication Disorders & Sciences

October 2001

A critique of markedness-based theories in Phonology

Naomi Gurevich

Eastern Illinois University, gurevicn@ipfw.edu

A CRITIQUE OF MARKEDNESS-BASED THEORIES IN PHONOLOGY

Naomi Gurevich
University of Illinois at Urbana-Champaign
ngurevic@uiuc.edu

The notion of *markedness* has been prevalent in phonology since its use by one of the founders of the Prague school of phonology, Trubetzkoy (1939). In contemporary writing it is most often used as a measure of the relative naturalness of linguistic elements. In this paper I explore the use of *markedness* in phonology literature and argue that it is an ill-defined notion that relies on circular reasoning and, quite often, leads to conflicting or vacuous predictions. Specifically, I question the generative theory-internal notion that markedness is encoded in the grammar. I focus on the multi-dimensional aspect of *markedness* in phonology and the various criteria used as diagnostics for the assignment of markedness values. I conclude that the predictive and descriptive powers of the dimensions of *markedness*, when taken individually, are far superior to those of *markedness* used as a cover term.

1. Literature Survey

1.1. Introduction

The goal of this paper is to explore the commonly accepted notion in phonology that markedness is somehow encoded in the grammar. I argue that, at best, markedness is a notion that exists only for linguists to make our job of describing languages easier and our generalizations more elegant. This argument is similar to the one made by Cole and Hualde (1998) against abstract Underlying Representations (UR's):

We find that, without any doubt, UR's are a valuable, and perhaps even essential tool for linguistic field work, whose adoption often allows for simple and concise statements of phonological patterns. However, our concern in this paper is not with the task of the linguist in identifying and describing sound patterns, but with the psychological status of UR's as an encoding [of] the sound representation of a word in the mental lexicon. (p.3)

Similarly, I question the psychological status of markedness as an encoding of some universal 'naturalness' in the phonology, although I readily admit that the notion of markedness may very well be a useful tool for linguists in a number of ways.

Providing conclusive evidence in favor of my argument will surely prove impossible. The existence of markedness as some psychological entity, much like Universal Grammar (UG) and many other linguistic mechanisms posited by generativist theorists, is an unfalsifiable hypothesis.[1] It is a *deus ex machina*: like the Greek literary mechanism used as a last resort to resolve overcomplicated plots, markedness is a supposed faculty that humans possess which can somehow resolve the tangled plot of phonology by linking the generativist-posited symbolic system to some physical reality. Despite the lack of evidence for the existence of some such faculty or even, as is shown in the present paper, agreement on what *markedness* means, it is a commonly used notion in the field of phonology as well as in other domains of linguistics.

Markedness goes by many names and a variety of definitions. It is a notion that is deeply embedded, although to varying degrees, in linguistic theories. And yet there is no consensus on what the term means. A similar sentiment is often articulated in other surveys that target the notion of markedness (recent work on this front includes Battistella 1990, 1996 and Rice 1999). Battistella (1990) notes that

> Different approaches to markedness (and there are many) define the markedness relation in different ways, apply the concept to different domains of inquiry, and integrate it into different theoretical approaches. (p.ix)

And Rice ends her paper with the following conclusion:

> Markedness is something about which linguists come to have strong intuitions. In many areas there is agreement: Something called markedness exists. It is multidimensional, with several factors involved at various levels (e.g., featural, combinatorial, positional). Variation exists in what can pattern as unmarked, although it is not without limit. However, many questions remain. […] The issues surrounding markedness do not appear to be ones that will find quick solutions, and the area promises to continue to be one of lively debate for some time to come. (p.37)

For the reasons detailed above, and others that will be expanded on within this paper, exploring the notion of markedness is both important and difficult.

In this part of the paper I consider and critique some of the major definitions of markedness that have been explored in the literature. I examine the entrance of the term *markedness* into mainstream linguistics jargon, how this notion gained ground, and how it became so closely tied to linguistic universals, and hence to

some psychological reality. I also attempt to isolate the often controversial theory-specific (more often than not, generative theory-specific) assumptions that are presupposed in various uses of markedness theory. The present discussion calls attention to the complications and misconceptions involved in the use of the notion of markedness in contemporary theories and helps build a case against the view that markedness is in some way significant to speakers.

My exploration of the notion of markedness in phonology literature is, at first, chronological. I start with the origin of the term in the writings of one of the founders of the Prague school of phonology, Trubetzkoy's *Principles of Phonology* (1939), where the notion of markedness is heavily utilized in the classification of types of neutralization. Chomsky and Halle note that "the notion of markedness is hardly mentioned in the phonological literature of the 1940s and 1950s" (1968:402). Hence *The Sound Pattern of English* (1968) is the next work I explore on the subject. Guitart's *Markedness and a Cuban Dialect of Spanish* (1976) draws on both Trubetzkoy's and Chomsky & Halle's work, as well as on Jakobson's (1968, 1971) and Postal's (1968), and offers a clear idea of the development of markedness after its entrance into the generativist framework. Herbert's *Universals and Markedness* (1986) and Greenberg's *Language Universals* (1966), as is obvious from the titles, further the view of markedness as a universal. Kaye's *On the Alleged Correlation of Markedness and Rule Function* (1979) showcases the complications of utilizing markedness notions within definitions of rule types in phonology. I finish with Rice's *Featural Markedness in Phonology: Variation* (1999), which begins with a review of various studies related to markedness and, perhaps inadvertently, leads to the conclusion that existing work complicates, rather than explicates, the situation.

1.2. Principles of Phonology (Trubetzkoy 1939)

It is commonly agreed that the term 'mark' originates with Trubetzkoy, one of the founders of the Prague school of phonology. The notion of markedness in his work is associated, to varying degrees, with articulatory complexity, combinatory possibilities of sounds, phonological statistics, functional load, and neutralization. Most prevalent, and most often cited, is Trubetzkoy's observation that the outcome of neutralization is normally the unmarked member of an opposition. Having borrowed the term from Trubetzkoy, generativists suggest that in contrast to their own approach he considered markedness assignments to be language-specific rather than universal; however, it is his work that sparked the use of markedness in the context of universal tendencies of language.

Trubetzkoy wrote that if two phonemes share the same set of features, except for one feature found in only one of the phonemes, this feature is the 'mark' and involves an extra articulatory gesture. Guitart (1976) points out that within the generative framework it is not always the case that the marked member involves an extra gesture. For example, a [+nasal] element is considered marked, but it is the unmarked oral sounds that require the extra gesture of closing the velum.

While the generativist argument that the marked element does not always require an extra gesture is well taken, the example shows how complex this situation really is. The resting position of the velum is neither completely open nor closed. Furthermore, in running speech the position of the velum depends on the surrounding sounds.

From a notational point of view, for Trubetzkoy the '+' member of an opposition usually corresponds to 'marked'. That is, [+nasal] is marked whereas [–nasal] is unmarked; similarly, [+glottal] is marked whereas [-glottal] is unmarked, etc. In adopting the term *markedness* generativists depart from Trubetzkoy's work in this respect as well. Since an element could be formally described as [+nasal] or [-oral], the '*marked = +*' aspect of Trubetzkoy's use of the term was abandoned.

In his classification of neutralization types, at the highest level of the typology Trubetzkoy distinguishes between contextually (in the environment of specific phonemes) and structurally (in specific positions in the word) conditioned neutralization. Within these two basic sets he also distinguishes between *dissimilative* and *assimilative* neutralizations.[2] The notion of markedness relates to the four types of dissimilative contextually conditioned neutralization. That is, these types of neutralization can serve as diagnostics for markedness values, as detailed below.

<u>Contextually Conditioned Dissimilative Neutralization</u>: The trigger for this type of neutralization is the same feature being neutralized (e.g. voicing triggers voicing neutralization); also, by definition, this involves change in polarity (+$\rightarrow$– or –$\rightarrow$+) when not privative. There are 4 possible types of this neutralization:

a) *<u>In the vicinity of both members of the same opposition</u>:* In many languages there is neutralization between voiced and voiceless obstruents in the vicinity of voiced and voiceless obstruents. (1) are examples from Serbo-Croatian:

    (1)   Serbo-Croatian obstruent voicing:
    a.   `srb`   'Serb. (masc.)'
           `srpski`   'Serbian'   `srpkinja` 'Serb. (fem.)'
    b.   `naručiti` 'to order'   `narudžba` 'the order'

In (a) the labial obstruent is voiced when it is word final (`srb`) but voiceless when it is followed by voiceless obstruents (`srpski` and `srpkinja`). Similarly, in (b) voicing of the palatal obstruent is determined by the voicing properties of the following obstruents: voiceless in `naručiti` but voiced when followed by *b* in `narudžba`. Another example comes from French, where the opposition between nasalized and nonnasalized vowels is neutralized before all (+/- nasal) vowels.

b) *<u>In the vicinity of the marked member, but retained in the vicinity of unmarked member</u>:* For example, in Slovak the opposition between long and short vowels

is neutralized after a syllable with a long nucleus. Another example comes from Sanskrit, where the opposition between dental and retroflex n is neutralized after retroflex ṣ (which occurs even with intervening vowels, labials, or gutturals), but is retained after nonretroflex consonants.

c) *In the vicinity of both members of a related phonological opposition:* For example, in Lezghian the opposition between rounded and unrounded consonants is neutralized before and after high vowels (u, ü, i) because these vowels are members of the opposition of timbre rounded/unrounded; but the low vowels (a, e) are not.

d) *In the vicinity of the marked member of a related opposition, but retained in the vicinity of the corresponding unmarked member:* For example, in Japanese, Lithuanian and Bulgarian the opposition between palatalized and nonpalatalized consonants is only phonologically valid before back vowels, and is neutralized before front vowels.

Types (a) and (c) cannot serve as diagnostics for markedness because the features in question could be either marked or unmarked; types (b) and (d) can, because there is a difference in results depending on whether the marked or unmarked feature is in the vicinity.

Trubetzkoy's work on Phonological Statistics, while not incorporated into his own notion of markedness, is relevant to later work on the subject. According to Trubetzkoy, statistics in phonology are significant because they can show how often a specific phonological element of a given language recurs in speech (token frequency), and more importantly for our purposes, they show the importance of the functional load of such an element or of a specific phonological opposition (type frequency):

By this method of examining the vocabulary it is also possible to determine for each language in numbers the extent to which the individual phonological oppositions are utilized distinctively (their functional load), as well as the average load of the phonemes in general. (pp.267-268)

However, markedness per se is never cited by the author as related to functional load. His discussion focuses on the relationship between neutralization and functional load, with markedness playing a role only due to its use in the definition of certain types of neutralization.

Trubetzkoy turns to Zipf's (1935) theory that frequency depends on the degree of articulatory complexity. Trubetzkoy rejects this theory, arguing that it is difficult to pin-point the degree of complexity of sounds; for example, which is more complex: tense vocal cords but relaxed organs of mouth, or lax cords and tense mouth organs? Trubetzkoy categorically rejects the explanation of phonological facts by means of biological, extra-linguistic, causes. However, he is

willing to translate Zipf's theory into phonological terms; for this he uses markedness. That is, markedness, rather than degree of complexity, determines frequency.

Such use of markedness as a substitute for some abstract notion, in this case the difficult to measure degree of complexity, suggests that markedness values are concrete, and even absolute. This tendency is echoed in the generative framework, where markedness values are linked to +/- feature values. And yet, when phonologists started work on determining markedness values they found themselves relying on many of the notions deemed abstract and in need of translation into markedness values, such as articulatory effort, for their decisions (e.g., Guitart 1976, Postal 1968, and Greenberg 1966 to name a few). This type of circular thinking can be observed throughout the literature on markedness.

On the subject of frequency, Trubetzkoy adds that one must take into account the possible presence and the extent of neutralization. That is, the differences between unmarked and marked opposition members, and the differences between oppositions that can be neutralized and those that cannot, affect phoneme frequency. Put simply, the presence and extent of neutralization affects the frequency of a sound: if neutralization can occur, the unmarked elements, which are the usual results of neutralization, will be more frequent. It becomes obvious that, in this diagnostic for frequency, it is not the markedness values that matter but whether or not the elements in question can be results of neutralization.

Trubetzkoy adds that this diagnostic does not always work because in some languages the markedness relationship cannot be objectively established. Since all other references to markedness in his work can be translated into 'presence of neutralization', this statement can be understood as the following: in some languages there may be cases of neutralization where no one member of an opposition is more likely to be the target of the process than the other member. In such cases, then, the presence and extent of neutralization does not affect the frequency of the elements in question. As discussed in §1.4, Guitart (1976) attempts to resolve the problem of neutralization where the outcome, contrary to expectations, appears to be marked. This becomes an important issue once, based on Trubetzkoy's work, the outcome of neutralization is formally declared a diagnostic for markedness values.

Frequency plays an important role in the "extent of utilization" of an opposition. By examining the vocabulary it is possible to determine for each language the extent to which the individual phonological oppositions are utilized distinctively, as well as the average load of the phonemes in general. Trubetzkoy distinguishes between "economical" and "wasteful" languages: the "economical" languages have numerous words that are only distinguished by one phoneme, and the percentage in which theoretically possible phoneme combinations are realized is very high. The "wasteful" languages have a tendency to distinguish words by

several phonological elements and to realize only a small percentage of the theoretically possible phoneme combinations.

Here Trubetzkoy discusses one of the key issues in phonology: the relation between sound substitutions and meaning. Neutralization will have different functional consequences for economical and wasteful languages. For economical languages the functional load of an opposition is expected to be higher, and phonological neutralization is more likely to yield functional neutralization.[3] It is vital to note that markedness has no functional purpose whatsoever. Its only connection to the subject is its use to distinguish between two of the types of contextually conditioned dissimilative neutralization. In Trubetzkoy's discussion of functional load, no one type of neutralization is singled out; in this particular discussion, markedness brings nothing to the party: it has no theoretical value.

In summary, Trubetzkoy not only brings the term markedness into the phonological arena; he also draws associations between this notion and many of the characteristics that markedness takes on in later literature (e.g., neutralization, functional load, and frequency), as well as its use as a translation mechanism from abstract notions such as articulatory complexity. However, as I suggest in the discussion of his work, in most cases Trubetzkoy intended to draw connections between the occurrence and result of neutralization processes, rather than markedness values per se.

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

In the final chapter of *The Sound Pattern of English* (SPE) Chomsky and Halle discuss unresolved problems from previous chapters. In particular, they acknowledge that their previous discussion suffers from the fundamental theoretical inadequacy of being overly formal and ignoring the fact that features have intrinsic content.

For example, their theory suggests that the "naturalness" of a class can be measured in terms of the number of features needed to define it. This diagnostic of naturalness often fails, as in the case of voiced obstruents: they have the more complex definition, as they are measured with a larger number of features than all voiced segments, but are intuitively more natural as a class than the class of all voiced segments (vowels as well as consonants). The authors point out that the content of the features, rather than the form of the definition, is responsible for naturalness. Hence, ignoring the content is precisely the reason for the failure detailed above.

Another example where their framework is found lacking is provided by independent phonological processes which somehow work together. E.g., Cole (1955) and Fudge (1967) subsume the following processes in Tswana under the single heading of "strengthening": voiced stops become ejectives after nasals, non-obstruent continuants become voiceless aspirated plosives, and obstruent

continuants become voiceless aspirated affricates. Chomsky and Halle agree, but lament the lack of a formal device to bring out the relation between the 3 processes.

A third example of where the framework is lacking centers on the fact that a vowel system as in (2) is more natural than those in (3) and (4), but the measures of evaluation previously suggested by the authors cannot make such a distinction.

(2) i   u   (3) i   u   (4) ü   ɨ
        e   o   e   o   ʌ
            a   æ   œ   a

Similarly, no distinction can be made between a more natural situation where all the vowels in a language are voiced, and an unattested situation where all vowels in a language are voiceless.

The above examples suggest a need to extend the theory in order to accommodate the effects of the content of features. This would allow the theory to distinguish between the natural and expected configurations and the less natural and less expected ones. In other words, the final chapter of *SPE* addresses situations where linguists can tell something is natural (whether it is one element as compared to another, a class of elements, a system within a language, or the relationship between some processes in a given language) but have no way to formalize this intuition within the framework as articulated at that point.

Note that these observations, used to motivate the need for a notion of markedness, are indeed valuable to the linguist; but the question of what value a language learner could derive from such knowledge, e.g., that some element in their language is more natural than another element, or that their vowel inventory is not as natural as their neighbors', is not addressed. This is due to the underlying generative belief (at least at that stage of the theory) that the job of the linguist is the same as the job of the learner, and that the theory's formalisms are equivalent to the manner in which learners code language. That is, that which is common across languages is presumed to be part of the UG and hence the knowledge learners are born with, and that which is unique about each language is what learners must acquire about their individual languages. The linguistic formalization of a language's grammar is, therefore, primarily a matter of identifying the exceptions and additional rules to the UG, which according to this theory mimics the job of a learner.[4] But while it makes sense for speakers of English to know that nasals assimilate in place of articulation to following stops in their language, what benefit is there to also being aware that this process is much more natural than, say, i→ɨ after non palatalized consonants (as occurs in Russian)? While arguing against this particular generative approach is beyond the scope of the present paper, I do try to emphasize that most knowledge encompassed by markedness issues is knowledge a learner either does not require (such as knowing a process in their language is more natural than a process in

some other language) or can easily determine with no prior knowledge (e.g., that intervocalically voiced stops are easier to articulate than voiceless ones).

An especially significant aspect of the framework detailed in *SPE* requires that rules that lead to the more natural configurations (e.g., voicing vowels, determining a vowel system as in (2), etc.) do not add to the complexity of a grammar because, as noted above, such rules are part of the UG and are therefore part of a speaker's knowledge already. Instead, the absence of such rules should increase complexity because a speaker must "unlearn" such rules if they are absent in their language. Thus, the Praguian notion of 'marked' and 'unmarked' values of features are incorporated into the framework in order to inject the intrinsic content of features into the theory; furthermore, the values are used in a manner such that unmarked values do not contribute to complexity.

> As a first step, the specifications in a matrix that constitutes a lexical entry are *u* (for unmarked) and *m* (for marked), along with + and –. We then add universal rules of interpretation which systematically replace the symbols *u* and *m* by the symbols + and –. Being universal, these rules are not part of a grammar but rather contentions for the interpretation of a grammar; they do not affect the complexity of a grammar, as determined by the evaluation measure, any more than the rules for interpreting → or { }. (pp.402-403)

Echoing Trubetzkoy's strategy of substituting markedness values for more abstract values, as if markedness values were any less abstract, *u* and *m* are offered as the content-wise replacements for the overly formal and content-lacking + and – . If intrinsic content of features could be evaluated in some absolute (non-abstract) manner, *u* and *m* might be considered less abstract than + and –. However, as noted above and repeated throughout my work, the diagnostics for assigning markedness values have remained abstract as well as controversial. In short, it is difficult to see how adding the abstract translation between phonetic reality and feature values can possibly resolve the problems of neglecting intrinsic content.

Finally, the authors suggest that with the use of markedness they now have the machinery for making distinctions between more and less plausible rules in purely formal terms.

> If it should prove possible to define a reasonably short list of such "plausible" phonological processes and show that all—or the majority of—the phonological processes encountered in different languages belong to this set, this would constitute a very strong empirical hypothesis about the nature of language. (p.427)

That is, what appears natural in human languages, once translated into markedness terminology, can be used to argue what is natural in languages. This type of circular reasoning is a recurring theme in markedness based theories.

1.4. Markedness and a Cuban Dialect of Spanish (Guitart 1976)

Guitart's work explores the benefits of the markedness version of generative phonology over the pre-markedness version. It is for this reason that his work is so useful in detailing the development of the notion of markedness after its incorporation into the generative framework. For Guitart, markedness can be divided into two theories: one is a more general theory of the phonological structure of human language; the second is a more specific theory of the phonological structure of the lexicon. I focus on the more general aspect in his work because the details of markedness and lexical representation, i.e., the intricacies of encoding information without complicating the grammar, are theory-specific and marginal to the present discussion.

In §1.3 I discuss the *SPE* notion that markedness helps formalize intuitions that linguists have based on the content of features. On a similar note, Guitart reiterates the idea that certain phenomena are somehow more natural in the languages of the world. For example,

1. Vowels can be voiceless, but are usually voiced; languages exist that have only voiced vowels, but there are no known languages that have only voiceless vowels.
2. When stops are neutralized in word-final position—a fairly common phenomenon—the result is usually the voiceless stop.
3. Commonly, languages have only two nasal consonants as systematic phonemes. When this is indeed the case, these two phonemes are most often *m* and *n*.
4. Children learning languages that have liquids usually master these sounds quite late, and before they do they tend to use glides instead.

These phenomena suggest that certain feature values are more common, more likely to appear in a given context, are acquired earlier than others, and are hence more natural. The already familiar strategy detailed in *SPE* is to characterize this naturalness, that is based on the intrinsic content of features, with markedness values. And, Guitart clarifies, since the set of features is universal, the intrinsic content of features is also universal. Hence if it is decided that it is more natural for back vowels to be rounded, that decision has been made for all languages.

Guitart is concerned with the heavy burden of making the universal marking decisions. He relies quite heavily on Postal (1968) for a discussion of marking criteria:

Relative occurrence of sounds: Certain sounds are very common in the languages of the world while others are rare. This leads to generalizations such as that nasality is marked for vowels because vowels are normally oral and there are no languages with only nasal vowels. Similarly, voicelessness is marked for vowels because no languages exist with only voiceless vowels.

Appearance in position of neutralization: This criterion is already familiar from Trubetzkoy's work where it originated as an observation rather than a diagnostic for markedness values. Postal does not rule out the possibility that for some features the marked value appears in a neutralized environment, so for him, this criterion is not absolute. Later in this section I discuss a strategy that Guitart suggests, which he terms "relative markedness", to deal with Postal's mistrust of this criterion.

Language acquisition and language loss: Guitart refers to Jakobson's (1968, 1971) hypothesis that sounds are acquired by the child—and lost by the aphasic—in a certain fixed, universal order. Those phonological elements that are acquired later and lost earlier are marked.

Phonological change and dialect variation: It should be expected that sound change would affect marked elements more often than the unmarked ones. This should lead to situations where marked features are lost; a hypothetical example is two dialects that differ in that one has both glottalized and plain consonants while the other has only non-glottalized consonants (which are cognate with both series in the former). Opposite situations, however, where there is a merger of unmarked to marked elements, would be extremely rare or even nonexistent.

Physiological and perceptual investigations: This, according to Postal, is the strongest evidence in favor of marking decisions. Physiologically speaking, features requiring special instructions to the organs of speech are expected to be marked. For example, apicality is marked for labials, palatals, and velars, but unmarked for dentals because this region is closest to the resting position of the tip of the tongue. On perceptual grounds, sounds that tend to enhance communication are regarded as unmarked, while marked sounds tend to obscure the signal. Note that according to this criterion clicks must be considered perceptually unmarked.

Implicational Hierarchy: The hypothesized order of acquisition underlies a fundamental assumption made by the theory of markedness, that the organization of features is hierarchical. That is, there are implicational relations among the sounds of a language. For example, the presence of fricatives in a language implies the existence of stops in the same language, because according to the theory, the feature 'continuant' is acquired after the feature 'obstruent'.

Syntagmatic assignments: With the exception of neutralization, the above markedness assignment criteria are based on paradigmatic properties. Syntagmatic markedness assignments can also be made on the basis of naturalness; that is, on the basis of what sequence of sounds are considered more natural. E.g., CVCV is a sequence that is most natural on both articulatory and perceptual levels.

At this point Guitart echoes my own criticism of markedness theory, discussed in §2, that the possibility the criteria may be in conflict is often ignored.

However, Guitart focuses his criticism on a much narrower scope than my own. He refers specifically to the situation where the physiological and perceptual criteria for markedness assignments conflict. He suggests as an example the case of voiced and voiceless obstruents: the voiceless ones are unmarked, but the voiced counterparts are chosen intervocalically because, according to him, while the voiced obstruents may be marked articulatorily, they are unmarked perceptually. Guitart does not discuss why intervocalic voicing is natural on perceptual grounds. While I do not doubt that a conflict between the physiological and perceptual criteria could arise, the case of obstruent voicing is not an example of such conflict, but instead, an example of the blurred line Guitart makes between syntagmatic and paradigmatic markedness assignments. That is, the markedness assignment of voicing in obstruents should be viewed as contextually determined. Intervocalic voicing of obstruents is generally agreed upon as unmarked articulatorily.

In any case, Guitart concludes that in situations where the physiological or perceptual criteria fail on their own, the use of both criteria at the same time may supply a fairly natural explanation. He points out the 'tug of war' aspect of the interaction between the two criteria, where certain feature values could be regarded as unmarked physiologically but marked perceptually, or vice versa. The author proposes the term "relative markedness" to describe the theoretical framework that would incorporate the following characterizations of phonological systems:

1. Not all sounds and sequences of sounds are either perceived or produced with equal ease.
2. The constraints imposed on the organization of phonological systems and on the utilization of phonological elements are due to the limitations of the human structures having to do with the production and perception of speech, and are, as such, universal.
3. Even though there are phonological elements that are both easy to produce and easy to perceive, there are elements that are physiologically simple but perceptually complex and elements that are physiologically complex but perceptually simple.
4. Neither maximal contrast nor least effort is the main guiding principle in human communication, i.e., neither the speaker nor the hearer is overwhelmingly preferred. Phonological elements which are neither too simple nor too complex—if judged by either the perceptual or the physiological criterion—will be used more frequently than elements which are least complex according to one criterion but most complex according to the other.
5. Human communication does not utilize sounds that are both harder to produce and harder to perceive.

The incorporation of these characterizations[5] into a theory leads Guitart to make the following formalization of three possibilities of markedness categories:

1. Elements that are unmarked both physiologically and perceptually
2. Elements that are unmarked physiologically but marked perceptually
3. Elements that are unmarked perceptually but marked physiologically
   ⋄ But no phonological elements that are marked both physiologically and perceptually.

We now see that "relative markedness" refers to a type of continuum on which every element is unmarked on some level, either physiologically, perceptually, or both. For Guitart, this means that the requirement that the output of neutralization be unmarked is always satisfied, and the reputation of the Trubetzkoy-inspired criterion is saved. He seems to have little concern that "relative markedness" allows any and all sounds to be unmarked on some level, rendering markedness values completely vacuous.

In summary, Guitart claims that markedness assignments are universal. When these assignments cannot be made based on physiological or perceptual grounds, they can rely on the relative occurrence and distribution of sounds in the languages of the world, including appearance in positions of neutralization; language acquisition and language disorders; and dialectal variation and sound change. For cases where the criteria fail to select the appropriate markedness value for some element, usually due to conflict between physiological and perceptual considerations, Guitart suggests that markedness can be relative, and proposes a formalization that renders any sound as unmarked on some level.

1.5. Universals and Markedness (Herbert 1986)

As part of his discussion of the underlying nature of prenasalized consonants and categorization of the types of half-nasal consonants among the world's languages, Herbert studies the relationship between diachronic and synchronic universals in phonology. Markedness surfaces in his discussion of universals.

Herbert describes a theory of markedness as articulated by Chomsky and Halle, and reiterated, for the most part, by Guitart. He cites a criticism by Lass (1972), that within markedness theory considerations of simplicity are not reconciled with language-internal economy of individual phonological systems. This criticism is based on the fact that, on a statistical level, the theory's claims about the frequency of some highly marked elements (e.g., front rounded vowels or clicks) might be justified; but as universals they do not hold for languages with the sound inventories of Swedish or Zulu. The author points out that Lass's interpretation of the theory, that the unmarked elements will always gain ground, is equivalent to claiming that all phonological evolution is directed towards the development of an optimal sound inventory. Markedness theory does not make such claims, and according to Herbert, it is able to offer some insight even for languages with highly marked inventories. That is, on a universal level, systems such as Swedish or Zulu can still obey some language-internal markedness considerations.

The author points out that the assignment of marked and unmarked values represents only a class of *observations*, but that statistical frequencies, diachronic mergers, synchronic neutralizations, etc. point to the fundamental correctness of this concept. However, "the theory remains one of observation, one of probabilities" (p.24). Herbert also concedes that the significance of the theory is weakened due to a total lack of explanation. He offers to provide some of this explanation in his classification of prenasalized consonants by pin-pointing and adding more phonetic information to some level of the symbolic system that links phonetic reality to feature values.

Herbert's response to criticism that markedness theory lacks explanation, adding more phonetic information to some level of this symbolic system, does not render the system any less symbolic or abstract. Herbert recognizes the fact that markedness theory is one of observations and implicational statistics. But the "fundamental correctness", as he puts it, of these observations nevertheless leads him, as it does most generativists, to suppose a universality that must be encoded in the UG.

1.6. Language Universals (Greenberg 1966)

Greenberg attempts to generalize the notion of markedness to mean some psychological reality, or even human nature, based on the fact that its characteristics can apply to phonology, grammar and the lexicon. He claims that the concept of markedness provides the possibility of finding more specific universals than can be arrived at by purely empirical methods. This rather grandiose statement is supported with circular arguments: Greenberg looks for the connection between markedness and universals by isolating the common features of markedness in the various fields of linguistics; he then uses the fact that he can identify such common features as evidence for the connection between markedness and universals.

Greenberg emphasizes the origin of the concept of markedness as arising from the Prague school of phonology in the context of problems of neutralization and the archiphoneme: although neutralization is a language-specific phenomenon, it is generally the same category of sounds which appears in the position of neutralization in the different languages, that is, the unmarked. He offers this as the first characteristic of markedness to be shared by other fields of linguistics.

The second characteristic of markedness that Greenberg cites, frequency, also originates in Trubetzkoy's work: the unmarked category has higher frequency (he does not specify *type* or *token* frequency, nor does he specify if within a language or universally). This is related to another characteristic of markedness, taken from Zipf's (1935) principle of least effort as well as Trubetzkoy's discussion of it, that the more complex is less frequent, and therefore, marked. A third characteristic of markedness is that the unmarked element will show greater

variety of subphonemic variation than its marked counterpart. That is, the more specified, the more marked, and the more variability (less resistance, more features can be filled) the less marked. A fourth is that the basic allophone is unmarked. E.g., in English [t] and [t^h] are allophones of /t/; [t] is the basic allophone, and is therefore the unmarked.

Having identified what he considers the main characteristics of markedness, Greenberg attempts to generalize them to Phonology, Grammar and the Lexicon. However, while the same major criteria of markedness seem to apply for grammar and lexicon, in phonology this concept is used differently. Greenberg's solution is to construct a grammar and a lexicon in terms of features. And so, going one-by-one through the above characteristics of markedness for phonology he "translates" them into the domains of grammar and lexicon. This translation of phonological characteristics of markedness into other domains then becomes Greenberg's evidence for the universality of markedness, and its existence as psychological reality, and even human nature. Certainly this is the best example of how markedness is part of the universals of linguistic theory, rather than of languages.

1.7. On the Alleged Correlation of Markedness and Rule Function (Kaye 1979)

Kaye's work was published in response to Houlihan and Iverson's *Functionally-Constrained Phonology*, appearing in the same volume. Specifically, Kaye presents data that calls into question the diagnostics proposed by these authors for whether a rule is neutralizing (contrast-obliterating) or allophonic (contrast-maintaining). Kaye's argument is based on the very formal definitions of neutralization and allophonic rules, which are, in part, dependent on markedness relations. The uncontroversial assumption that a rule cannot be both neutralizing and allophonic is complicated by the expectation that a rule that starts out diachronically as either allophonic or neutralizing must remain as such over time in the grammar.[6] Most importantly for the present paper, Kaye's work emphasizes the circular nature of markedness definitions, as well as definitions of other phenomena that formally depend on markedness (e.g., neutralization).

Houlihan and Iverson state that rules which convert marked segments into unmarked ones must be neutralization rules, and all rules which are not neutralizing are allophonic. Given this definition, Kaye points out that it is hardly surprising that neutralization results in less-marked segments. Furthermore, since the presence of marked segments in a language typically implies the presence of their unmarked counterparts (we know this as implicational hierarchy), it is to be expected that rules that result in unmarked segments will do so in the context of a phonological inventory that already contains these segments. In the same line of reasoning, the claim that rules resulting in relatively marked segments are allophonic rules follows from the fact that marked segments are found less frequently and are thus less likely to be found among the underlying phonemes of

a language or as the output of a phonological rule applying before the rule in question.

Kaye concludes, therefore, that the fact that these definitions are *generally* true is in no way an argument in their favor, but merely a consequence of their formalizations. Put simply, defining these two types of rules using markedness relations, when markedness relations are basically defined using the rules they participate in, is circular and provides no insight into either situation.

To illustrate his point Kaye discusses an example from Algonquin where a rule converts marked segments to unmarked ones, but is obviously an allophonic—and not neutralization—rule. This, and other examples in Kaye's work, clarify how Houlihan and Iverson's principles fail because their definitions lead to contradictions. In many dialects of Algonquin there is a rule that devoices initial obstruents. This initial devoicing rule turns relatively marked segments (voiced) into relatively unmarked ones (voiceless) and ought to be a neutralization rule following Houlihan and Iverson's definitions. But in fact it is an allophonic rule because although both voiced and voiceless obstruents exist in Algonquin, according to Kaye only voiced obstruents occur in word-initial position—the context in which initial devoicing applies. Thus it is clearly a meaning-maintaining, and not a meaning-obliterating, rule. This example shows that basing definitions of rule types on markedness rather than meaning considerations is ineffective.

In summary, the author illustrates violations of Houlihan and Iverson's definitions, and emphasizes that the fact their principles generally hold is beside the point. This fact follows from general notions of markedness as well as the circular definitions of neutralization and allophonic rules, and therefore plays no role in constraining the class of possible phonologies.

1.8. Featural Markedness in Phonology: Variation (Rice 1999)

Rice's work comprises an extensive survey of literature on markedness in phonology. Rice puts together a nearly exhaustive list of the various definitions, used in sources such as Jakobson, Trubetzkoy, Kenstowicz, and other works on phonological theory and writings on the theory of markedness, to describe the difference between marked and unmarked categories. This list of definitions is reproduced in (5).

(5)

Marked	**Unmarked**
Less natural	More natural
More complex	Simpler
More specific	More general
Less common	More common
Unexpected	Expected
Not basic	Basic

Less stable	Stable
Appear in few grammars	Appear in more grammars
Later in language acquisition	Earlier in language acquisition
Subject to neutralization	Neutralization targets
Early loss in language deficit	Late loss in language deficit
Implies unmarked feature	Implied by marked feature
Harder to articulate	Easier to articulate
Perceptually more salient	Perceptually less salient

This list, surely, represents the intuitions most linguists have about elements that are characterized using markedness terminology. These are the general observations that can be made about such elements; the general tendencies that are often repeated in elements belonging to the same category (of marked or unmarked), and hence considered universal.

Rice's own approach to markedness, which stems from the study of language acquisition, is to treat it as related to structure: markedness is mostly a consequence of the amount of structure, where the less structure, the less marked and vice-versa. Her goal is no different from similar works that stem from other disciplines within linguistics. It is to provide "an account of what universal grammar allows to be unmarked and what the universal and language particular aspects to markedness are" (p.34).

Rice and Avery (1995, cited in Rice) argue that two aspects of language must be accounted for: that there is cross-linguistic uniformity in the features that pattern phonologically as unmarked ("global uniformity"), but these features are not always the same ("local variability").[7] The variability aspect is, of course, what makes the pattern only a *tendency* rather than a *law*. For Rice, it seems, the key to explaining markedness resides in the criteria that separate elements that succumb to the general tendency to pattern with others of the same markedness category, from those that do not.

Rice devotes the bulk of her study to reviewing existing work on markedness, adding her own approach at the end. However, she does not appear to be able to make any generalizations that encompass the works she reviews, save for recognizing the fact that markedness is a controversial area that will continue to be debated for some time to come.

1.9. Discussion

The one constant that is retained, and reiterated, in the sections above relates to the origin of the notion of markedness as a universal. This is the observation that unmarked elements tend to pattern similarly cross linguistically, as well as to some extent within languages. Encoding markedness categories in the grammar is deeply rooted in the recognition that such universal tendencies exist. This is

emphasized in generative theories, where universal tendencies are especially significant because these are the aspects of languages that are attributed to the UG.

The trend of universalizing markedness statements starts with the observation, credited to Trubetzkoy, that there is a similarity between the elements that can appear in positions of neutralization across languages. Regardless of Trubetzkoy's original intention for using the notion of 'mark', which was to help classify types of neutralization, this generalized the language-specific processes of neutralization. For linguists, who are interested in generalizations, this finding is not one that can be ignored. Thus, the notion of markedness from its very beginning embodies some universal aspect of a theory of language.

The trend of using markedness as a translation system for more abstract or symbolic notions, a trend which mistakenly regards markedness values as less abstract or symbolic than the notions it translates, can also be traced back to Trubetzkoy. Trubetzkoy used markedness to translate the extra-linguistic biological causes of articulatory complexity into a linguistic system. Chomsky and Halle use markedness as a link between phonetic reality (that is, content of features) and the purely symbolic system of +/- feature values, creating linking rules that are somehow encoded in the grammar without complicating it. And Greenberg finds a novel use for markedness as a translation system: he uses it to translate phonological characteristics into other domains of linguistics.

Steriade (1995) outlines exactly how markedness statements, in the form of redundancy rules that may be left unspecified lexically, might ease the burden of the learner by moving this burden into the UG. A typical problem with the generativist approach arises: in an attempt to simplify the task of the learner, as much of the burden as possible is moved into the UG; the tradeoff aspect of this theory—the fact that this supposed entity acquires more and more complexity—is conveniently ignored because the UG can be anything that is required of it; it is not a physical entity that can ever be examined.

Whether one believes in the UG or not, it is still interesting to ask the following questions: What is the advantage of including markedness statements in the grammar? Why should properties of language that are easily observable (such as ease of articulation[8]), and those that are not at all important to learners of a language (such as cross-linguistic frequency of some element) be built into the linguistic mechanism we are supposedly born with? Never mind the questionable advantage of removing the burden from being the responsibility of a learner, what is the advantage of possessing such knowledge at all? Why should a learner have access to the knowledge that sonorants are voiced before any exposure to the data, when the minimum of such exposure would make this fact clear? It seems obvious that any and all advantages of encoding markedness statements are theory-internal: the only motivation for including markedness statements in the grammar is that such statements represent some universal tendencies. The goal of generative

phonologies is closely tied to isolating such tendencies for the express purpose of building a UG.

Beyond questioning generativist wisdom, in this section I emphasize the circularity of so many of the markedness definitions and uses. First and foremost is the circularity of using markedness values to characterize aspects of language that are difficult to measure (e.g., articulatory effort, frequency), which are in turn used as criteria for awarding markedness values. Second, there is Greenberg's search for the connection between markedness and universals by isolating the common features of markedness in various fields of linguistics, which are in turn used as evidence for the existence of this connection. Third comes Kaye's illustration of how defining rule types as either neutralizing or allophonic using markedness values is not insightful because markedness values themselves depend on the types of rules an element can undergo: markedness values are awarded to elements based on the behavior of these elements within a system, and participation in neutralization or allophonic rules is what makes up this behavior.

Finally, this section shows that, while few agree on exactly what markedness means, nobody questions its encoding in the grammar. The generativist-internal motivation for the existence of such a component in the grammar is clear: the UG contains (but is not limited to) that which is common across languages; markedness captures these commonalities, and must therefore be part of the UG. External motivation for allowing markedness to represent some psychological reality is nonexistent, and yet such an essential part of scientific reasoning does not seem lamented in the literature.

2. The Multi-Diagnostic Approach

In this section I draw on the various characterizations of *markedness* reviewed in the previous section in order to show motivation for, but also problems with, the multi-diagnostic approach to markedness in phonology.

2.1. Motivation for the multi-diagnostic approach

The discussion in §1 of the history of the term *markedness* and the development of this notion in linguistics suggests that this notion was never uni-dimensional. Trubetzkoy's (1939) use of the term is associated most commonly with neutralization processes, that is with the fact that the output of such processes is usually the unmarked member of an opposition. But his work also links markedness to articulatory effort, where the marked member of an opposition is defined as requiring an extra articulatory gesture as a means of explaining the biological causes of such effort in linguistic terms.

In bringing the term into the generative framework, and thereby into contemporary phonology theories, Chomsky and Halle (1968) attempt to formalize some degree of naturalness based on the intrinsic content of features. While

naturalness could be considered the one leading dimension of markedness, it is clear that it is a dimension that is inherently abstract and cannot be determined based on any one diagnostic. The most notable reasons for this are that naturalness is based on articulatory as well as perceptual grounds, which don't always agree, and that it is context dependent rather than absolute. Furthermore, not only is naturalness in itself a multidimensional measure but its use in the generative framework is applied to a variety of phenomena which are common among the languages of the world (listed in §1.4).

Within the generative framework, markedness values can be awarded to elements as compared with each other within an inventory system, a class of elements, an inventory of elements from one language as compared to another language, the relationship between some processes in a given language, and processes among those possible in all languages. Since the naturalness of so many different aspects of language is measured using markedness, it is obvious that no one diagnostic could accommodate the full load (e.g., what diagnostic could possibly measure both frequency and voicing distinction?). In short, the fact that markedness values accommodate a variety of phenomena is uncontroversial in phonology literature, and the multi-diagnostic determination of markedness values is necessary in order to oblige the variety, and could therefore also be considered uncontroversial.

For a list of diagnostics for assigning markedness values, I turn to Postal (1968). He discusses SPE's framework of allowing marked-unmarked to represent + and – feature values, and notes:

> Accepting such a theory involves the responsibility for discovering the right class of universal rules interpreting M and U representations as + and –. This is a vast undertaking. At the moment, from the point of view of a completed system, our knowledge along these lines is limited. But there is already a great deal of knowledge, and many M-U decisions can be made with some confidence. (p.168)

The diagnostics he suggests may be used with confidence are the following (these are cited in §1.4, where they are discussed in detail):

A. Relative occurrence of sounds
B. Appearance in position of neutralization
C. Language acquisition and language loss
D. Phonological change and dialect variation
E. Physiological and perceptual investigations
F. Implicational hierarchy
G. Syntagmatic assignments

An additional, and relevant, checklist for assessing markedness can be found in McMahon's (1994) discussion of natural morphology, where naturalness is defined in terms of markedness:

> Unmarked or natural features occur frequently cross-linguistically; appear often in numerous contexts in languages where they occur; are relatively resistant to change but often result from changes; occur in pidgins and are introduced early in Creoles; and are acquired early by children, but unaffected or lost late in aphasia. Furthermore, borrowings and neologisms in a language will typically follow the unmarked pattern; and it is rarely affected by speech errors, although marked forms are commonly assimilated to the unmarked pattern in error. (p.98)

The motivation behind the multi-diagnostic approach to markedness is summarized as follows: Markedness is meant to encode some degree of what linguists agree is 'intuitively natural'. What is considered natural by linguists is normally based on phenomena that recur or are common in many of the world's languages. For generativists this implies a connection to the UG (i.e., what is common between languages is usually part of the UG, and should not add complexity to the learning of any individual language). For non-generativists some degree of naturalness is also a useful measure as a general observation that can be made about human languages (that is, there is an intellectual benefit to making generalizations regardless of whether one is building a UG or not). The characteristics that are shared by many languages are in themselves multi-dimensional, and range from articulatory complexity, to behavior in certain processes, to order of acquisition. For these reasons markedness value assignments necessarily involve a variety of diagnostics.
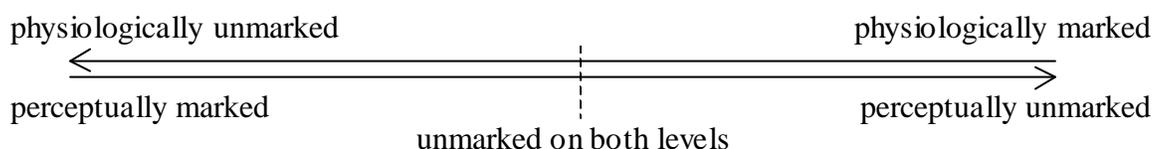
2.2. Problems with the multi-diagnostic approach

There is an inherent problem with assigning only one of two values (marked or unmarked) based on what, as explained above, must necessarily be a variety of, sometimes unrelated, criteria. A choice of one of only two values is far too limiting and cannot be expected to resolve cases where the various dimensions involved in assigning markedness values are in conflict. Thus, when the diagnostics for markedness contradict each other the predictive and explanatory value of markedness is weakened.

In addition to the problem of conflicting markedness value assignments, there is a misconception that markedness drives linguistic processes (e.g., that it is the markedness status of some element that renders it more, or less, stable diachronically). Since markedness is used to evaluate a variety of phenomena that are common among the world's languages, it emerges as a cover term for the multitude of characteristics of language it can represent. A close examination of the individual facets of language subsumed under *markedness* illustrates the conflicting, and even vacuous, predictions they make. This leads to one important

conclusion: it is not markedness per se that drives linguistic behavior but the processes that are subsumed under the concept of markedness.

Guitart's 'relative markedness', discussed in §1.4, constitutes the clearest example of a case where the multi-diagnostic determination of markedness assignments leads to vacuous predictions. Postal voices reservations about the use of neutralization as an almost absolute criterion in assigning markedness values, a practice that can be traced back to Trubetzkoy's work. Postal points out that the output of this process is not necessarily the expected unmarked member of an opposition. In response, Guitart suggests his 'relative markedness' theory, that—although this does not seem to bother the author—basically allows any element to be unmarked on some level of the scale between physiological and perceptual markedness (see Figure 1).

physiologically unmarked   physiologically marked

⇐――――――――――――――――――――⋮――――――――――――――――――――⇒

perceptually marked   perceptually unmarked

unmarked on both levels

Figure 1

Guitart suggests there is a 'tug of war' between the physiological and perceptual criteria: an element can be unmarked on both dimensions (i.e., fall in the middle of the scale in Figure 1), or it can be unmarked on one dimension or the other, but never both. The illustration helps clarify that an element cannot be marked on both physiological and perceptual grounds (because it cannot be on both edges of this continuum simultaneously), and that an element could fall anywhere on the continuum and be considered unmarked ('relatively unmarked') on some level. In this manner the result of neutralization can always be unmarked, and hence, can always serve as a criterion for markedness assignment, despite Postal's reservations. However, this scale renders every element unmarked on some level, and hence renders the physiological and perceptual diagnostics of markedness assignments, the diagnostics deemed most important by Postal and others, vacuous.

Markedness value assignments based on both syntagmatic and paradigmatic criteria constitute examples of the contradictory predictions that can result from the multi-diagnostic aspect of such assignments. Consider the simple case of voicing: on a syntagmatic level, voicing intervocalically is aerodynamically natural, while on a paradigmatic level voicing inventory-wise is marked (on the basis of articulation effort; also on the basis of implicational hierarchy: if a language has voiced stops it has voiceless ones; along other criteria as well, e.g., frequency) (Westbury and Keating 1986). While most phonologists readily acknowledge that naturalness is contextually dependent, few consider the contradictory predictions of syntagmatic vs. paradigmatic criteria for the determination of markedness values a defect of the theory.

In Rice's (1999) discussion of the patterning of the unmarked she focuses on three phenomena that are common: emergence of the unmarked, submergence of the unmarked, and transparency of the unmarked. Emergence of the unmarked refers to the fact that unmarked elements, often the 'default' elements in their class, are usually the choice for epenthetic matter. An example is the epenthetic i in Cairene and Iraqi dialects of Arabic:

(6)　Epenthetic i breaks up triconsonantal clusters (Ito 1989:241-242)
　　　a. CCC → CCiC e.g., /ʔul-t-l-u/ → ʔultilu 'I said to him' (Cairene)
　　　b. CCC → CiCC e.g., /gil-t-l-a/ → gilitla   'I said to him' (Iraqi)

Submergence of the unmarked refers to a very different behavior of unmarked elements. Those elements which are most susceptible to assimilation, deletion, and other forms of weakening are diagnosed as unmarked. E.g., the coronals in Korean, where the coronal feature is a target, assimilating to other places of articulation (a & b) but not a trigger (c & d), and is therefore considered unmarked:

(7)　Korean place of articulation assimilation (Cho 1988, Iverson & Kim 1987, cited in Rice 1999:8)
　　　a.　coronal-labial　　ko/tp/alo　　ko[pp]alo　'straight'
　　　b.　coronal-dorsal　　pa/tk/o　　　pa[kk]o　　'to receive and'
　　　c.　labial-coronal　　pa/pt/o　　　pa[pt]o　　'rice also'
　　　d.　dorsal-coronal　　ka/ŋt/o　　　ka[ŋd]o　　'rubber'

Unmarked features also display a common behavior that Rice calls transparency of the unmarked, where they are transparent to assimilation while the marked features block it:

Steriade (1987) shows that vowel harmony may be restricted to cross laryngeals but not other places of articulation, suggesting that laryngeals are unmarked in place while other consonantal places of articulation are marked; Paradis and Prunet (1989) argue that vowel harmony crosses the coronal place of articulation in Guere, but not others, implying that coronal is unmarked amongst the other places of articulation. (Rice 1999:9)

These three ways in which unmarked elements tend to pattern similarly cover a wide range of phenomena, some exact opposites of each other. What theoretical significance is there to a diagnostic of markedness that is based on participation in processes such as loss as well as epenthesis, and assimilation as well as transparency to it? As with 'relative markedness', although to a lesser extent, almost any element can be characterized as unmarked based on its participation in one of so many processes.

A final example where the multi-diagnostic determination of markedness values fails comes from diachronic change. In particular, I explore the

contradiction between the predictions markedness values make about the rate of change: marked elements are usually considered less stable diachronically, but often they are regarded as the more stable elements.

Marked elements are harder to learn and are less natural and hence are often subject to weakening or loss. This view of the marked element as being less stable diachronically has already been presented in §1.4 where the example of glottalized consonants disappearing from an inventory is offered. Another example of diachronic change of marked elements towards their unmarked counterparts is the intervocalic weakening of stops:

(8)   Lenition/weakening (Hock 1991:81):
        Lat. pacatum > (*)pagado > Sp. [paɣaðo]
        > dial. [paɣaØo] 'pacified, pleased'

In this weakening of intervocalic stops from Latin to a dialect of Spanish, first the intervocalic stops are voiced in the intermediate, reconstructed, stage. Then they are spirantized in Spanish, and the dental element is finally lost in the dialect of Spanish. Since each stage of the change in (8) involves a more "relaxed" pronunciation than the previous stage and hence a less marked element, the weakening process is a change of marked elements to unmarked ones.

However, marked elements are often more stable diachronically. Consider an example from morphology: irregular verbs in English. The more common irregular verbs resist regularization into *–ed* forms (e.g., *drink/drank/drunk*, *eat/ate/eaten*, *is/was/were*, etc.). There are even cases where the irregular form (considered marked) takes over an unmarked form, as in the case of *dove* replacing *dived* as past for *dive*. The example is taken from Hock (1991:175):

(9)   Old English   New English
       dȳvan   dive
       dȳv(e)de   dived → dove   (cf. drive : drove)[9]

Thus certain irregular verbs in English are more stable, less susceptible to change, diachronically. Of course, one could claim that their frequency—the most common explanation for their stability—renders them unmarked, and hence their stability is expected. The question becomes one of determining the markedness values of these elements in a situation where more than one value is predicted by the diagnostics, and hence one of conflicting predictions.

An example of diachronically stable marked elements in phonology comes from click-languages. Most would agree that clicks are marked inventory-wise in almost every aspect, but few would predict the loss of these sounds from the inventory of Zulu, for example. Also, glottalization, offered above as an example of a marked feature that can be expected to be lost, is actually quite common in syllable-final stops (Silverman, pc).

Since marked elements can be both more and less stable diachronically, it is doubtful that the markedness status of such elements is behind their behavior. In fact, it becomes obvious that a variety of phenomena, and the interaction between them, must be responsible for the variation in diachronic stability between these elements (e.g., functional load, frequency, contrast maintenance, and social factors, among others). These are the same phenomena used as diagnostics for the determination of markedness value assignments. Subsuming them under one cover term, and representing them with only one of two values, ignores not only the specifics of these phenomena and their influence on diachronic stability, but also the interaction between them, and hence reduces the explanatory power of the theory.

3. Conclusion

In §1 the exploration of markedness in phonology literature concludes that the term is an abstract notion whose definitions are mostly circular and whose motivation is limited to theory-internal reasoning. In §2 the circularity of the previous section's definitions is echoed in the conclusion that the multi-dimensional aspect of markedness is, in part, a paradox. On the one hand, the fact that markedness is used to evaluate a multitude of often unrelated phenomena in a variety of domains (within a class, an inventory of one language, or all the world's languages) makes the determination of markedness values necessarily multi-diagnostic. On the other hand, the various criteria used as diagnostics often lead to conflicting assignments of markedness values.

But there is a way out: I suggest that the phenomena subsumed under the cover term *markedness* are individually far more valuable than the cover term itself. It is these phenomena that often drive phonological behavior. The fact that these phenomena often conflict with each other, and therefore lead to conflicting or vacuous predictions, is only a problem if they are thought of as dimensions of one aspect of language, that is markedness. Individually, the fact that these phenomena don't always act in unison, as if towards some predetermined goal, is part of the nature of language. Furthermore, funneling the generalizations that can be made based on the individual phenomena into one of only two values (marked or unmarked) emerges as a generalization that is too wide in scope to be truly insightful. The predictive and descriptive powers of the dimensions of markedness, when taken individually, are therefore far superior to those of markedness used as a cover term.

NOTES

1 Again I turn to Cole and Hualde's (1998) questioning the existence of abstract underlying representations. They believe that "the burden of the proof should be on the defenders of non-observable entities" (p.8). I agree, but do not believe that the defenders of markedness as some psychological entity can be persuaded to take on this burden.

2 In both cases he means to/from the "contextual phoneme" in a slightly different use of the terms than is common in contemporary writings. Trubetzkoy classifies as *dissimilative neutralization* processes where the trigger of neutralization is the same feature being neutralized (e.g., voicing triggers voicing neutralization). *Assimilative neutralization* refers to processes where an independent feature triggers loss of contrast (e.g., nasals trigger voicing neutralization in obstruents).

3 *Functional neutralization* refers to the consequence for meaning that a phonological neutralization of a distinctive opposition may have. That is, if a distinction between two sounds is obliterated, it may lead to the obliteration of the meaning distinction between lexical elements that contain these sounds (and may previously have been distinguished by them).

4 This is an oversimplification of the theory, especially since the UG is not limited to commonalities across languages. My intention is only to show the theory-internal relationship between the job of a linguist and that of a learner.

5 Guitart uses *perception* somewhat inaccurately, neglecting the fact that perception is significantly shaped by language specific contrasts. For example, for speakers of English, where no such contrast exists, it is difficult to perceive the difference between glottalized and plain stops; but for Korean speakers it is easy.

6 As far as I know, there is no reason for this to be true.

7 It is not clear to me what aspect of this notion is *local*. Possibly, she means within a language as opposed to the *global* cross-linguistic tendencies. Or, more likely, she means *locally* within the group of elements that pattern as unmarked cross-linguistically. I can only guess, but it does not really matter for our purposes.

8 This property is difficult to measure objectively; but it is not objective measuring that I refer to here.

9 This type of change is a 4-part analogy, a systematic type of proportional analogy that generalizes a pattern of morphological relationship between given forms to other forms.

REFERENCES

BATTISTELLA, Edwin L. 1990. *Markedness: The Evaluative Superstructure of Language*. State University of New York Press: Albany, NY.

——. 1996. *The Logic of Markedness*. Oxford University Press: New York, NY.

CHOMSKY Noam and Morris HALLE. 1968. *The Sound Pattern of English*. Harper & Row Publishers: New York, Evanston, and London.

COLE, Desmond Thorne. 1955. *An Introduction to Tswana Grammar*. London.

COLE, Jennifer S. and José I. HUALDE. 1998. *The Object of Lexical Acquisition: A UR-free Model*. Proceedings of CLS 34.

DINNSEN, Daniel A. (Ed.). 1979. *Current Approaches to Phonological Theory*. Indiana University Press: Bloomington & London.

GLEASON, Jean Berko (Ed.). 1993. *The Development of Language*. Macmillan Publishing Company: New York, NY.

GOLDSMITH, John A. (Ed.). 1995. *The Handbook of Phonological Theory*. Blackwell Publishers: Cambridge, MA.

GREENBERG, Joseph H. 1963. *Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements*. In Greenberg, Joseph H. (Ed.). "Universals of Language". Second Edition. MIT Press: Cambridge MA.

——. (Ed.). 1963. *Universals of Language*. Second Edition. MIT Press: Cambridge MA.

——. 1966. *Language Universals, With Special reference to Feature Hierarchies*. Mouton & Company: The Hague, The Netherlands.

GUITART, Jorge M. 1976. *Markedness and a Cuban Dialect of Spanish*. Georgetown University Press: Washington DC.

HERBERT, Robert K. 1986. "Language Universals, Markedness Theory, and Natural Phonetic Process". *Trends in Linguistics, Studies and Monographs 25*. Winter Werner (Ed.). Mouton de Gruyter.

HOCK, Hans Henrich. 1991. *Principles of Historical Linguistics*. 2nd Ed., Revised and Updated. Mouton de Gruyter: Berlin and New York.

HOCK, Hans Henrich and Brian D. JOSEPH. 1996. *Language History, Language Change and Language Relationship: an Introduction to Historical and Comparative Linguistics*. Mouton de Gruyter: Berlin and New York.

HOULIHAN, Kathleen and Gregory K. IVERSON. 1979. *Functionally-Constrained Phonology*. In Dinnsen, D. A. (Ed.). "Current Approaches to Phonological Theory". Indiana University Press: Bloomington & London.

INGEMANN, Frances (Ed.). 1976. *1975 Mid-America Linguistics Conference Papers*. The Linguistics Department, University of Kansas: Lawrence, Kansas.

ITO, Junko. 1989. "A Prosodic Theory of Epenthesis". *Natural Language and Linguistic Theory* 7: 217-259.

JAKOBSON, Roman. 1963. *Implications of Language Universals for Linguistics*. In Greenberg, Joseph H. (Ed.). "Universals of Language". Second Edition. MIT Press: Cambridge MA.

JESPERSEN, Otto. 1964. *Language: Its Nature, Development and Origin*. W.W. Norton & Company: New York.

KAYE, Jonathan Derek. 1979. *On the Alleged Correlation of Markedness and Rule Function*. In Dinnsen, D. A. (Ed.). "Current Approaches to Phonological Theory". Indiana University Press: Bloomington & London.

KENSTOWICZ, Michael and Charles KISSEBERTH. 1979. *Generative Phonology*. Academic Press, Inc.: San Diego, CA.

LABOV, William. 1994. *Principles of Linguistic Change*. Blackwell: Oxford, UK and Cambridge, MA.

MCMAHON, April M. S. 1994. *Understanding Language Change*. Cambridge University Press: Great Britain.

PARADIS, Carole and Jean-François PRUNET (Eds.). 1991. *The Special Status of Coronals: Internal and External Evidence*. Phonetics and Phonology, Volume 2. Academic Press, Inc.: San Diego, CA.

POSTAL, Paul M. 1968. *Aspects of Phonological Theory*. Harper and Row: New York, Evanston, and London.

RICE, Keren. 1999. *Featural Markedness in Phonology: Variation*. Part 1. GLOT 4.7:3-6. Part 2. GLOT 4.8:3-7.

RICE, K. and P. Avery. 1995. "Variability in a deterministic model of language acquisition: a Theory of Segmental Elaboration". *Phonological Acquisition and Phonological Theory*. Archibald, J. (Ed.). 23-42. Lawrence Erlbaum: Hillsdale, NJ.

SMITH, Neilson V. 1981. "Consistency, Markedness and Language Change: On the Notion 'Consistent Language'". *Journal of Linguistics* 17: 39-55. Cambridge University Press: London.

STERIADE, Donca. 1995. *Underspecification and Markedness*. In Goldsmith, John A. (Ed.). "The Handbook of Phonological Theory". Blackwell Publishers: Cambridge, MA.

TRUBETZKOY, N.S. 1939. *Principles of Phonology*. Translated by Christiane A. M. Baltaxe. 1971. LA: University of California Press.

VENNEMANN, Theo. 1969. *Sound Change and Markedness Theory: On the History of the German Consonant System*. In Stockwell, Robert P. and Ronald K. S. Macaulay (Eds.). "Linguistic Change and Generative Theory. Essays from the UCLA Conference on Historical Linguistics in the Perspective of Transformational Theory". Indiana University Press: Bloomington and London.

——. 1988. *Preference Laws for Syllable Structure and the Explanation of Sound Change With Special Reference to German, Germanic, Italian, and Latin*. Mouton de Gruyter: Berlin.

WAUGH, Linda R. 1979. *Remarks on Markedness*. In Dinnsen, D. A. (Ed.). "Current Approaches to Phonological Theory". Indiana University Press: Bloomington & London.

WESTBURY, J. R. and P. A. KEATING. 1986. "On the Naturalness of Stop Consonant Voicing". *Journal of Linguistics* 22: 145-166.
ZIPF, G. K. 1935. *The Psycho-Biology of Language*. Houghton Mifflin: Boston.