On a universal criterion of rule coherence

Michael Shapiro

One of the factors commonly retarding progress in linguistic theory is a kind of amnesia affecting its exponents, whereby well-grounded, highly productive principles of language structure discovered in the past are forgotten in contemporary discussions. The principle of *isomorphism* has, alas, suffered from just such neglect — with predictable consequences.[1] In its simplest form this principle states that different levels of language structure embody identical rules of organization. The first explicit application of this principle was made in Jakobson (1932), and the first significant recognition of the pervasiveness of isomorphism between the different levels of language was achieved by Hjelmslev (1938).

One of the consequences of the discovery of the isomorphism principle was an important shift in the understanding of linguistic arbitrariness (associated chiefly with de Saussure). Due largely to Jakobson's studies after the war, it became increasingly clear that the core of language is constituted by extensive patterns of similarity and difference among the shapes of grammatical morphemes which correspond to relations of similarity and difference among their meanings. In semiotic terms such correspondences between relations on the expression level and the content level of languages are to be understood as *diagrams*, relations represented by relations. In recent discussions (to the limited extent that this idea is explored in depth)[2] mapping relations of this sort have come under the designation of "iconicity" in grammar (e.g., Haiman 1985).

There is another, equally fundamental sense in which isomorphism can be said to pervade the structure of language, namely the sense in which rules at the core of grammar are not merely statements of regularities but are *coherent*. The notions associated with the terms "rule" and "coherence" need to be examined separately. Although the concept of rule was not prominent among the theoretical advances of the early European structuralists, it is nonetheless clear that its ubiquitousness today owes much to an understanding of grammatical relations as patterning and regularity that goes back to prewar discussions (principally

in Prague and Copenhagen) of the foundations of linguistic theory. What is missing from both pre- and postwar theorizing, however, is the notion of the coherence of linguistic relations, and as a corollary, the precise means whereby coherence is to be expressed in the practice of linguistic description.

All along, the potential for making coherence an explicit principle in the understanding of language structure existed unexploited among the many overt achievements of early structuralism, specifically in the idea of *markedness*. Coherence obtains when rule relations signify the mirroring of markedness values, either parallelistically or chiastically (more on this later), across content and expression levels, or between different aspects of expression (as in the case of some morphophonemic congruences). In the latter case — which will be the focus of this paper — the term *automorphism* has been proposed (Haiman 1985: 4, cf. Shapiro 1987: 165), imported into linguistics from mathematics.[3] Since patterning is present at all levels of grammar, to the extent that the rules of language structure expressing this patterning reflect congruences of markedness values we can attribute their coherence (their *raison d'être*) to such cohesions. What is more, we can do this uniformly by virtue of the isomorphism of grammar. Nothing proves the validity of this universal notion of coherence better than the evidence of linguistic change. The drift of a language involves the actualization of patterns that are coherent in just this sense and the rejection of those that are not (Shapiro 1985; cf. Plank 1979, 1981; Andersen 1980b: 203).

Rules are more than mere generalized formulas of patterns when they embody specifications of coherence between linguistic elements, namely cohesions between units and contexts. This criterion of rule coherence remains true and valid but practically vague without the necessary involvement of markedness because it is markedness that provides the explicit means of expressing coherence. In order to illustrate how this criterion works to make linguistic analysis aspire to explanatory adequacy I have chosen a set of morphophonemic alternations in English that affect the distribution of tenuis and media obstruents as between nouns (also a few adjectives) and verbs:

	Nominals		Verbals
[-f]	*proof*	[-v]	*prove*
	safe		*save*
	wife		*wive*

[-θ]	*mouth*	[-ð]	*mouth*
	sheath		*sheathe*
	loath		*loathe*
[-s]	*diffuse*	[-z]	*diffuse*
	close (adj)		*close*
	refuse		*refuse*
	advice		*advise*
	(ab)use		*(ab)use*
	excuse		*excuse*
	house		*house*

This alternation between /f θ s/ and /v ð z/ also affects (albeit to a much lesser extent) English plural formation, e.g., *staff/staves* [obsolete], *hoof/hooves, wolf/wolves, oath/oaths* (cf. *cloth/clothes,* identifiable only by etymological analysis), and *house/houses.* Since English orthography regularly reflects the media obstruent in the case of [-v] by writing *v* and [-ð] by writing *e* after *th* (not always; cf. *sheathe* but *mouth* above), the interesting and numerically greater set of examples is the one exhibiting an alternation between [s] and [z] — which subsumes orthographically atypical items like *advice* and *device.* The alternation is basically unproductive as it affects noun/verb pairs in contemporary English, but its persistence requires ultimately that some accounting be reached as to its *raison d'être,* which is to attempt an answer to the question "Why does it persist?" It is certainly not unusual for languages to rationalize unproductive pockets of the grammar by investing morphophonemic alternations with just the sort of motivation we are seeking here (cf. Shapiro 1969). Where no motivation arises, an alternation is liable to be eliminated, as it is in some samples of the English case under consideration (see below).

To proceed, then, to answer the question in terms of grammatical coherence, one must examine the markedness values of the sounds and the meanings of the grammatical categories in which they are implemented. Taking the markedness values of the meanings first, we know that verbs are marked *vis-à-vis* nouns in that verbs necessarily make reference to the time axis while nouns do not. Given the marked value of verbs and the correspondingly unmarked value of nouns we would expect the distribution of tenuis and media obstruents to mirror the markedness values of the grammatical categories.

At this point we have to be careful about exactly what "mirror" means. In most discussions of iconicity or isomorphism, it is taken for granted

that the mapping of sound/meaning relations is one of simple replication; indeed, where the values are at variance with each other, such instances are termed "counter-iconic" (Wurzel 1984: 204).[4] My own extensive investigations of this issue (with particular reference to the structure of Russian; cf. n. 3) have led me to the conclusion that replication of markedness values between expression and content is typically confined to morphology, defined as the domain of relations between basic signs and contrasted to morphophonemics, defined in turn as the domain of relations between contextual variants of the same linguistic sign(s) (Andersen 1969: 807; cf. 1980b: 89–90). In privileging chiastic semiosis – complementation of markedness values – morphophonemics is only being true to its nature, i.e., as the domain of grammar defined by contextual variation of signs (= alternation). Because this is an area of the theory of grammar that has not been sufficiently investigated, any assertions must necessarily be tentative. Thus, chiasmus as a typical form of semiosis in morphophonemics may ultimately turn out to pertain to stem alternation but not to the contextual variants of desinences (cf. Andersen 1980a: 43–44).[5]

To a significant extent, of course, this whole way of looking at linguistic coherence turns on the determination of the markedness values themselves – not always an easy or straight-forward task for learners as well as analysts. But there is a heuristic element to which the notion of coherence gives learner and linguist alike direct access, and it is this element that becomes particularly helpful in resolving problematic cases (Shapiro 1974b; cf. Anttila 1980: 276 for an assessment). In the English data utilized for purposes of illustration above, for instance, it is clear that tenuis obstruents co-occur with the unmarked category of nouns, while media obstruents co-occur with the marked category of verbs. Assuming this to be an instance of morphophonemic iconicity affecting stem alternants and invoking the principle of chiastic semiosis, we ought to expect the markedness relations between sounds and meanings to be oppositely valued (= markedness complementarity). We have already stipulated the values of the two categories involved, noun and verb, but what about the stem-final obstruents?

Trubetzkoy claimed (1962: 142) that it is impossible to say whether English has distinctive voicing or distinctive protensity (tense vs. lax) in its obstruent system, but the concept of rule coherence with its reliance on markedness considerations now makes this agnosticism seem groundless. Positions of neutralization are diagnostic in this respect because neutralization rules provide contexts in which variation rules tend typi-

cally to produce diagrams of the markedness values of the terms of phonological oppositions (Andersen 1979: 381). The relation between syllable peaks and contiguous obstruents in English is such that syllable peaks are [−long] before tense obstruents but are [+long] before lax obstruents, sonorants, and in final position. Hence *beet* is [bit], but *bead, beam,* and *bee* are [biːd], [biːm], and [biː], respectively. Beyond the fact that the complementary distribution here is semiotically significant − as a sign of the non-distinctiveness of quantity in English − it is the stipulation of the tenuis and media obstruents as distinctively tense vs. lax that allows the variation rule to be coherent. Since tense obstruents are marked relative to the unmarked lax obstruents, and shorter realizations understood as abridgements of syllable peaks are marked relative to unabridged peaks (which are unmarked), the markedness values of the vowels replicate those of the contiguous obstruents. If the tenuis and media obstruents were assumed to implement the opposition voiced vs. voiceless (as they often erroneously are),[6] the variation would lack coherence because the markedness values of the obstruents would not match those of the vowels (voiced obstruents are marked, voiceless unmarked).

This little digression into the phonological problematics of English is necessary, apart from its practical bearing on the argument about rule coherence, because it illustrates the methodological status of what is now commonly referred to as "independent motivation" in linguistic explanation. Trubetzkoy's perception of irresolvability of the English tenuis/ media problem in its obstruent system can be seen as justified only as long as circularity is barred from explanations of language structure (as if language were not a hermeneutic object). The mutual dependency of the elements of the solution proposed above − the shorter realizations of the syllable peaks seen as abridgments rather than the longer ones as prolongations, the stipulation of protensity as the relevant phonological category rather than voicing, and the invocation of markedness considerations as the vehicle of grammatical coherence − testifies to the fact that these elements cohere as an ensemble of conditions informing the data. This is the structural coherence that emanates from an evaluation (intrinsically, in the grammar) of the units and the contexts in tandem, in a mutually dependent manner, so that randomness and arbitrariness are reduced to a minimum (if not always to nil).

In an unavoidably circuitous way we now come back to the earlier examples of alternation between English tenuis and media obstruents in stem-final position. The alternation becomes coherent when we understand it to subsume two conditions: (1) the alternation is morphopho-

nemic, hence the markedness values as between sounds and meanings
will be chiastic (complementary); (2) the obstruents involved are distinc-
tively tense vs. lax. The recognition of these two conditions enables us
to assert a coherence based on markedness values. Otherwise the alter-
nation would be strictly arbitrary, non-iconic, and non-coherent.

When units and contexts do not cohere, the typical outcome in the
long run is a heightened tendency toward the reduction of such instances,
to the limit of their wholesale elimination from the language. With
reference to our English examples there are attested historical changes
that confirm the correctness of the analysis, specifically by showing
morphophonemic coherence in just the sense advanced to be the telos of
the changes. Where coherence has already been reached, no further
changes occur. Thus in the history of English there is evidence (Horn
1954: 799 – 802) of generalization of either the tense or the lax obstruent
in words which now regularly have orthographic *s*, e.g., *enterprise, com-*
promise, purpose, promise, practise (= American *practice*). In Middle
English texts one can observe the testing of the contemporary rule in the
occasional writing of *z* instead of *s*, particularly in verbs (but not only).
In a Milton manuscript, for example, one finds the spelling *practiz'd*; and
as late as 1836, the pronunciation of the infinitive with a *z* is proscribed
as vulgar by normative grammarians (Horn 1954: 800). The same differ-
ence in pronunciation as between the nominal and the verbal forms of
the word evidently obtained for *enterprise* and *compromise* in Middle
English, the difference here being in the particular obstruent that was
generalized. In the case of *practise* and *promise*, it was the tenuis obstruent
that was generalized; in *enterprise* and *compromise* it was the media
obstruent. Exactly why it was *s* in the first pair and *z* in the second
constitutes a separate problem that might be treated in the spirit of rule
coherence and markedness, but I hesitate to offer an explanation. Perhaps,
in the presence of a primarily or secondarily stressed vowel in the verb
form there is a discernible tendency to generalize the unmarked media *z*
(*compromise*, but also *close*, as in *at the close of* ...). Similarly, one might
want to explain the gradual elimination of the pronunciation of *greasy*
with [-z-] as a case of an unmotivated alternation being dropped from
the (standard) language: there is no chiastic distribution of markedness
values within the category of nominals. The semantic split between *louse*
and *lousy* would tend to confirm such an analysis albeit obliquely.

Although the focus of this exposition has been on phonology and
morphophonemics, there is no doubt that the principle of isomorphism
in grammar would predict the equal and uniform applicability of the very

same notion of coherence to an understanding of lexis and syntax. Some efforts in that direction are available here and there in Haiman (1985), but Battistella (1985) is the first study known to me that utilizes my methodological strategy exactly in explaining a set of related syntactic changes. The contemporary English use of the semantically empty or periphrastic *do* that serves as a carrier of tense in interrogative, negative, and emphatic sentences lacking an auxiliary verb (*Did anyone tell you the news?; I didn't get the notice; I did read the assignment!*) is absent from Old English. It can be explained as the result of a development wherein the telos of change is determined by an alignment of markedness values (gradually, over the span from Old English to Modern English) that is generally analogous to the morphophonemic example, even though the specific form of semiosis (parallelism rather than chiasmus) is different. The markedness-based isomorphism in these syntactic examples accounts not only for the spread of *do* from its originally non-periphrastic locus to semantically marked paradigms or contexts but also explains its loss in the unmarked simple declarative forms.[7]

There may be ways of defining linguistic coherence other than by appealing to markedness isomorphism. But coherence thus construed is clearly a theoretical advance in the understanding of grammatical structure and language change as unitary phenomenological domains, and where structure at any given synchronic point is always the cumulative teleological result of change. The passage of some five decades since the heyday of our Prague and Copenhagen predecessors need not be an insuperable obstacle to present-day efforts at achieving a potent theory of grammar. More cooperation unfettered by ideological or socio-academic biases is called for (cf. Itkonen 1986). In this enterprise the work of naturalist grammarians in the German-speaking countries of Europe, of North American exponents of "natural" or functional syntax, and of linguists with a semiotic outlook on both sides of the Atlantic will need to emphasize common structuralist foundations and points of conceptual convergence in order that progress be made on all methodological fronts.[8]

Notes

1. I am using the term "isomorphism" in the traditional sense to refer to grammatical relations and not in the sense explored by writers like Givón (1985) or Haiman (1985), Haiman (ed., 1985), namely as a way of characterizing (*grosso modo*) the diagrammatic relation between language and world or grammar and reality. Givón and Haiman refer to such relations as "meta-principles" and talk about the way in which a linguistic code

encapsulates experience. As Itkonen (1986: 40) points out, this "discourse-pragmatic" sense of isomorphism (and, for that matter, of iconicity) goes back at least as far as the modistic grammarians. Note that Haiman (esp. in the edited volume (1985), but also before and since) wants to distinguish between isomorphism and diagrammatization, on the one hand, and isomorphism and economy, on the other, for which he has been appropriately criticized by Givón (1985: 188) and Itkonen (1986: 39), respectively.

2. See especially Dressler (1985: Chap. 10) and the references therein; also Wurzel (1984, 1985); Mayerthaler (1981; now available in English translation), and Plank (1979, 1981). A compact English source for work on morphology along naturalist lines is now represented by Dressler (et al. 1987). All of these studies are by scholars who have been influenced to one extent or an other by Jakobson (esp. 1965). Their narrow understanding of the concept of icon and iconicity has been justly criticized by Andersen (1980b: 202).

3. In a rare concession of priority, Andersen (1980a: 34) testifies that "patterns... where one morphophonemic alternation mirrors another so that they together form what may be called an automorphic structure in the paradigm, were first described explicitly by Shapiro, who gives several examples of existing or emerging patterns of this kind from contemporary Russian (1969)." Since Plank, Wurzel, and Mayerthaler do not cite my work and habitually take their exploitation of the idea of iconicity back to Jakobson through Anttila, this is perhaps the appropriate time to set matters straight. Anttila (1969) cites my paper "Toward the recognition of iconicity in language" which was given in December 1968 at the Winter Meeting of the Linguistic Society of America in New York and then became part of Shapiro (1969); (he also cites the manuscript version of Shapiro 1971). Unfortunately, Anttila (1969) appeared in Finnish and was thus not digested outside of Finland; however, Anttila (1972) refers to Shapiro (1969) and the semiotic ideas promulgated therein. In subsequent publications, Anttila is scrupulously consistent in giving credit to his former UCLA colleague, as well as to Andersen. In a series of publications following on my 1969 book, I continue to credit Andersen (1966) and his succeeding publications (see esp. the long lists in Shapiro 1976 and 1983). Anttila (1977: 4) gives a brief historical view of the same period that overlaps to some extent with the one I have just given. With the prominent exception of Dressler (who has exchanged offprints with me over many years), the members of the Natural Morphology school who most often recur to iconicity (Wurzel and Mayerthaler) never cite Shapiro (1969) or subsequent publications (like Shapiro 1974a, an important milestone in my thinking). I trust that the contacts initiated immediately before and during the meetings in Krems will help overcome these lacunae in future.

4. Andersen (1980b: 202) points out that Wurzel (as Greenberg before him, on whom he relies) goes wrong here because he fails to take the special relationship between genitive plural and nominative singular desinences in Russian into consideration. I could add that Wurzel seems totally to ignore the idea of markedness complementarity (as developed, for instance, in Shapiro 1974a, and recapitulated in Shapiro 1980 and 1983).

5. Battistella (1990: Chap. 3) explores this matter in some depth, but the question still remains open.

6. One is reminded here of the persistence, in American linguistic literature at least, of the mistaken notion that standard German has distinctive voicing in its obstruent system (cf. Andersen 1972: 45, n. 22). The careful analysis in Vachek (1976) supports the claim that English has distinctive protensity and not voicing.

7. Battistella (1990) recapitulates his earlier analysis of these constructions and embeds it in a much wider framework with many further examples.

8. For an extended attempt to cross-pollinate naturalist ideas about grammar with Peircean semiotic, see Dressler (1985: Chap. 10); cf. also Dressler (et al. 1987).

References

Andersen, Henning
1966   Tenues and mediae in the Slavic languages: A historical investigation [Unpublished Ph. D. Thesis, Harvard University.]
1969   "A study in diachronic morphophonemics: the Ukrainian prefixes", *Language* 45: 807 – 830.
1972   "Diphthongization", *Language* 48: 11 – 50.
1979   "Phonology as semiotic", in: S. Chatman et al. (eds.) *A semiotic landscape* (The Hague: Mouton), 377 – 381.
1980a   "Morphological change: towards a typology", in: J. Fisiak (ed.), *Recent developments in historical morphology* (The Hague: Mouton), 1 – 50.
1980b   "Summarizing discussion: Introduction", in: T. Thrane et al. (ed.) *Typology and genetics of language* (= *Travaux du Cercle Linguistique de Copenhague* 20) (Copenhagen: Linguistic Circle of Copenhagen), 197 – 210.
Anttila, Raimo
1969   *Uusimman äännehistorian suunnasta ja luonteesta* [About the trends and character of the newest historical phonology] *(= Publications of the Phonetics Department of the University of Turku 5).*
1972   *An introduction to historical and comparative linguistics.* (New York: Macmillan).
1977   *Analogy.* (The Hague: Mouton).
1980   "Language and the semiotics of perception", in: I. Rauch – G. F. Carr (eds.) *The signifying animal: The grammar of language and experience* (Bloomington: Indiana University Press), 263 – 283.
Battistella, Edwin
1985   "Markedness isomorphism as a goal of language change: the spread of periphrastic *do* in English", *Lingua* 65: 307 – 322.
1990   *Markedness: The evaluative superstructure of language.* Albany: SUNY Press.
Dressler, Wolfgang U.
1985   *Morphonology: the dynamics of derivation* (Ann Arbor, Mich.: Karoma Press).
1987   "Semiotische Grundlagen einer Theorie der Natürlichen Phonologie und Morphologie", in: *Semiotica Austriaca,* herausgeg. v. J. Bernard (Wien: Österreichische Gesellschaft für Semiotik), 165 – 172.
Dressler, Wolfgang U. – Mayerthaler, Willy – Panagl, Oswald – Wurzel, Wolfgang U.
1987   *Leitmotifs in natural morphology.* (Amsterdam: Benjamins).
Givón, Talmy
1985   "Iconicity, isomorphism, and non-arbitrary coding in syntax", in: J. Haiman (ed.), 187 – 219.
Haiman, John
1985   *Natural syntax.* (Cambridge: Cambridge University Press).
Haiman, John (ed.)
1985   *Iconicity in syntax.* (Amsterdam: Benjamins).
Hjelmslev, Louis
1938 [1970] "Essai d'une théorie des morphèmes", *Essais linguistiques.*[2] (København: Nordisk Sprog- og Kulturforlag), 152 – 164.
Horn, Wilhelm
1954   *Laut und Leben: Englische Lautgeschichte der neueren Zeit (1400 – 1950),* II, edited by M. Lehnert (Berlin: Verlag der Wissenschaften).

Itkonen, Esa
     1986   "Form-meaning isomorphism, or iconicity, in diachronic linguistics (and
                    elsewhere)", in: M. Remmel (ed.) *Symposium on formalization in historical
                    linguistics: Summaries,* (Tallinn: Academy of Sciences of the Estonian SSR),
                    38–46.
Jakobson, Roman
     1932   "Zur Struktur des russischen Verbums", in: *Charisteria Guilelmo Mathesio
                    oblata* (Praha: Pražský Lingvistický Kroužek), 74–84 [reprinted in Jakob-
                    son 1971: 3–15].
     1965   "Quest for the essence of language", *Diogène* 51: 21–37 [reprinted in
                    Jakobson 1971: 345–359].
     1971   *Selected writings,* II: *Word and language.* (The Hague: Mouton).
Mayerthaler, Willi
     1981   *Morphologische Natürlichkeit.* (Wiesbaden: Athenaion).
     1988   *Morphological naturalness* [translated by Janice Seidler]. (Ann Arbor, Mich.:
                    Karoma Press).
Plank, Frans
     1979   "Ikonisierung und De-Ikonisierung als Prinzip des Sprachwandels", *Sprach-
                    wissenschaft* 4: 121–158.
     1981   *Morphologische (Ir-) Regularitäten.* (Tübingen: Gunter Narr).
Shapiro, Michael
     1969   *Aspects of Russian morphology: A semiotic investigation.* (Cambridge, Mass.:
                    Slavica).
     1971   "Markedness and Russian stress", *Linguistics* 72: 61–77.
     1974a   "Morphophonemics as semiotic", *Acta Linguistica Hafniensia* 15: 29–49.
     1974b   "Tenues and mediae in Japanese: A reinterpretation", *Lingua* 33: 101–114.
     1976   *Asymmetry: An inquiry into the linguistic structure of poetry.* (Amsterdam:
                    North-Holland).
     1980   "Russian conjugation: theory and hermeneutic", *Language* 56: 67–93.
     1983   *The sense of grammar: Language as semeiotic* (Bloomington: Indiana Uni-
                    versity Press).
     1985   "Teleology, semeiosis, and linguistic change", *Diachronica* 2: 1–34.
     1987   "Sapir's concept of drift in semiotic perspective", *Semiotica* 67: 159–171.
Trubetzkoy, Nikolai S.
     1962   *Grundzüge der Phonologie* ³(Göttingen: Vandenhoeck & Ruprecht).
Vachek, Josef
     1976   "The opposition of voice and tension in modern English paired consonants",
                    in: J. Vachek, *Selected writings in English and general linguistics,* (Praha:
                    Academia), 364–372.
Wurzel, Wolfgang U.
     1984   *Flexionsmorphologie und Natürlichkeit: Ein Beitrag zur morphologischen
                    Theoriebildung* (= *Studia Grammatica* 21) (Berlin: Akademie).
     1985   "Zur Determiniertheit morphologischer Erscheinungen: Ein Zwischenbe-
                    richt", *Acta Linguistica Academiae Scientiarum Hungaricae* 35: 151–168.