MARKEDNESS, CAUSATION, AND LINGUISTIC CHANGE: A SEMIOTIC PERSPECTIVE

MICHAEL SHAPIRO
Brown University

1. *Instead of prolegomena: a philosopher's-eye view of language*

Philosophers have not been much interested in the material aspect of language, apart from the fact that there must be one. In fact, one could assert that philosophers have not begun to approach the real phenomenon of language.[1] But some of what linguists say may sound naive to a philosopher. Perhaps a bridging of the gap between the two disciplines can be essayed here by way of approaching the several points I would like to make about markedness, causation, and linguistic change.

Philosophers have been exercised about the ontological status of meaning. When a linguist (of the semiotic-structuralist persuasion) makes signans and signatum (the material and the intelligible part, respectively, in Jakobson's gloss) the two parts of a sign, this seems to be treating them as being ontologically on a par. But to a philosopher, the signatum is, though equally real, of a very different ontological type. The signans would remain the same material thing even if it had no significance (although then it would probably not have been produced). But the signatum is only a potentiality consisting in the interpretability of the signans. That is why a philosopher would want to say that the signans *is* the sign—i.e. what is interpretable—though it is a sign (and hence a signans) only because it *is* interpretable (nonarbitrarily). The signatum, then, is (using Peirce's terminology) the immediate object of the sign or signans, answering to the immediate interpretant.

---

[1]This essay's specific formulation of the issues (like Shapiro 1991) owes much to my correspondence with T. L. Short (beginning in 1981) and to a careful study of his writings on Peirce, especially the *semeiotic* (see References). I am also grateful to Henning Andersen for having invited me to participate as a discussant in his Workshop on Actualization Patterns in Linguistic Change and to contribute to this volume, as well as for his encouragement during the several stages leading up to its publication. I am, of course, solely responsible for all and any shortcomings.

That is why, rightly or wrongly, linguists' talk of 'content system' and 'expression system' (à la Andersen 1984) makes a philosopher of language uneasy. While it is true that the material aspect of language tends to develop into a diagrammatization of its meaning, yet the meaning is not *there* (as either a Platonic object or as thoughts or concepts in the minds of language users) independently of the 'expression system' or system of signantia. There is no meaning or content except *in* the interpretability of the material signantia. There is no discursive thought, no concepts except in our learned capacity to use the material signantia—whether in production for others to interpret or in interpreting the linguistic acts of others, or in that internalized speaking to oneself that we call 'thinking'. That, at least, is the doctrine of such diverse philosophers as Plato, Wilfrid Sellars, and Peirce. So there is not a content system independent of the expression system, on this philosophy of language: there is only a shared set of rules for (1) forming expressions and (2) interpreting them. And when expressions are interpreted they are not replaced (e.g., in the mind of the interpreter) by content. For there is no 'content' that can stand by itself in that way. Instead, on Peirce's view, decoding is translation, i.e. from one material signans to another (whether in the same or a different expression system) or from material signans to actions, feelings, or habits that are not themselves signs in the same sense. Now these actions, feelings, and habits cannot themselves be what is meant by 'content system'. The content system must be, instead, *either* the rules of interpretation by which the translation is made *or* the immediate objects represented by the original signs and grasped or apprehended in this interpretation or translation of these signs (into new signs or into such ultimate interpretants as habits, actions, or feelings).

Philosophers of language do not ordinarily mention phonemes or even morphemes. They tend to begin with complete words, sentences, arguments (Peirce's semes, phemes, and delomes). Aristotle says in his *De Interpretatione* (16a1): "First we must define the terms 'noun' and 'verb', then the terms 'denial' and 'affirmation', then 'proposition' and 'sentence'. (In the next work, the *Prior Analytics*, he goes on to treat arguments.) Continuing (16a20): "By a noun I mean a sound significant by convention, which has no reference to time, and of which no part is significant apart from the rest". And (16b6): "A verb is that which, in addition to its proper meaning, carries with it a notion of time. No part of it has any independent meaning, and it is a sign of something said of something else". Finally (16b26): "A sentence is a significant portion of speech, some parts of which have an independent meaning ...." Now Jakobson and Halle (1971) define morphemes "as the ultimate constituents endowed with proper meaning" and as "the smallest semantic vehicles". But a

morpheme might be only part of a word, since, e.g. in English, *un-* and *-ly* as prefix and suffix, respectively, contribute the same thing to the meaning of the various words to which they are affixed, though they never mean anything if left to stand alone. Philosophers of language follow Aristotle in making the word their basic or simplest unit of analysis, even though it is not clear that Aristotle's definition of nouns and verbs really distinguishes them from morphemes that are not words. What does Aristotle mean by "no part is significant apart from the rest?" For *un- has* a significance, albeit no replica signifies anything by itself (except in metalinguistic utterances like "All the *un*s are the same"). If we return to what Aristotle says about meaning, we get no help (16a3): "Spoken words are the symbols of mental experience and written words are the symbols of spoken words. Just as all men have not the same writing, so all men have not the same speech sounds, but the mental experiences, which these directly symbolize, are the same for all, as also are the things of which our experiences are the images". Why does *ungainly* symbolize experience and not *-ly?* Of course, there is no *ly*-ness found except as part of *ungainliness*, *loveliness*, etc., but neither is any *ungainliness* found except as part of *ungainly boys, ungainly horses*, etc.

The problem of how to draw the line between word meaning and morpheme meaning (in the case of those morphemes that are less than whole words) involves the more general problem of determining what meaning is. In the last passage quoted from Aristotle he appears to have no notion that language might be a determinant of experience. That is mostly true, being only slightly mitigated by what he says about concept formation (e.g., at *Posterior Analytics* II, 19). But putting that problem aside, he clearly locates meaning 'in the head', or, more accurately, in individuals' experiences of the world. Why not in the world directly? Or why not in neither, but in some realm of abstract entities—if not Plato's timeless forms, then in Popper's "World Three" of cultural artifacts? In any case, philosophers since Aristotle have been much exercised about the ontology of meanings. And one may wonder whether, on that point, even despite their failure to distinguish morpheme meaning from word meaning, philosophers might not well be in advance of linguists.

For example, when linguists conventionally speak of language as a "bridge between meaning and sound" and identify meaning with the "nonlinguistic real or imagined world, the things we talk about", this is still cruder than Aristotle. Perhaps for much of linguistics it does not matter that we have no clear idea of what meaning is: as long as we can express the meanings of sentences, words, and morphemes in *other* words—as long as such translations are available—we can express points about which things mean what, which mean the same, which are different ways of meaning the same

thing, what meaning a morpheme contributes to the meaning of a word, etc. We can make such points without ever saying what meaning is or whether there are any such things as meanings at all.[2] But a general theory of language would seem to require a clear conception of meaning—or whatever it is that language, to be more than mere sound, is supposed to 'express'.

Linguists like Hjelmslev (following Saussure) define language as mediating "between two realms of substance—the things about which we speak and the physical tokens" and then proceed to distinguish expression and content. So content would appear to be *the things we talk about*, or what would otherwise be called 'meaning'. Now we might speak of Napoleon, of trees, of this piece of paper, so content would seem to include physical things and persons. For this reason a philosopher may be confused when a linguist says that language must comprise "two distinct sign systems, a system of content signs and a system of diacritic signs". Since content is not 'in' language at all, except for that part of it that makes up the sounds speakers make in speaking, by 'content sign' a linguist cannot mean content. Diacritic signs, then, are other aspects of language. But is there a real difference? The smallest units of the content system are morphemes, and the diacritic signs are what Jakobson calls 'distinctive features'. But morphemes are made up (of phonemes that are in turn made up) of distinctive features. Why call these two distinct sign systems, rather than signs (morphemes) and their material parts (distinctive features)?

The signantia of the diacritic signs are the familiar Jakobsonian distinctive feature terms, and these signs all have the same signatum, viz. otherness. But to a philosopher this is tantamount to saying that their sole function is to be different from one another. And is that, in itself, a semiotic function? To be such as to be recognized as "this is that and not the other" is not the same as *signifying* otherness. A complementary problem affects the content system, since content signs are said to have, for the most part, no signantia of their own apart from those of the diacritic system. Which is to say

[2]To say as Peirce sometimes said that meaning *is* translation is to say that there is no such *thing* as meaning. As Wilfrid Sellars puts it: to say that German *rot means* red is simply to say that from a certain point of view *rot* and 'red' are to be classed together. To say they *have* the same meaning does *not* mean that there is some third thing (their meaning) that they have: it means only that they belong to the same class of words. But what defines these classes is the function or use of their members in the representative languages of those members. Hence Wittgenstein's "meaning is use". But all of this is too nominalistic unless we understand, with Peirce, that use or function presupposes the reality of laws, rules, habits. Use and function also implicate the future. Hence it is possible to identify meaning also with either rules or future effects. As Peirce says, the reality of a rule is not exhausted by the present: its reality consists in its influence on actual events, including future events. The meaning of a legisign has to be explicated in terms of general rules or habits, but the meanings of its replicas in terms of future effects, actual or potential.

that the expression system constitutes the sign vehicles of the content system: the content signs *are* syntagms of diacritic signs (phonemes). And diacritic signs themselves are not really signs at all: they are only the vehicles of morphemes and, hence, they reduce to the distinctive features that constitute them. Instead of two systems of signs we have a distinction between a system of material constituents of signs and the signs formed out of these materials. The attempt to make these into two separate systems *of signs* leads linguists to invent signs without significance (diacritic signs) and signs without identity (content signs).

The reason linguists divide language into these "two basic subsystems" is that they *are* two distinct systems, albeit not two systems *of signs*. As is shown clearly in the case of nonsense words, a native speaker recognizes phonemes of his own language even when they signify nothing, whereas of a language he does not know, he cannot distinguish phonemes from inarticulate noises or, at best, he cannot identify different occurrences of the same phonemes. So there is a system at that level and then another governing possible combinations and the interpretation of those phonemic syntagms that constitute morphemes, words, etc. Both systems together constitute the semiotic phenomenon of language, but there is no advantage (much less necessity) to viewing each separately as a system of signs.

Now suppose we maintain that morphemes are signs distinct from the distinctive features that constitute them (which sounds like a flat contradiction). Then how are morphemes to be identified? If one says that content signs form oppositions, which make up the system of meanings in language, strictly on the basis of their signata, this cannot be to say that we distinguish one morpheme or word from another because of what they severally mean. We distinguish them from one another by the distinctive features in which they differ. Whether they mean the same or different things depends on the rules of their interpretation (i.e. on the habits of interpretation common to the individuals of the community that speaks that language: the interpretants such habits determine may be emotional, energetic, or logical). In light of this, a philosopher would have some difficulty understanding a linguist's assimilation of encoding and decoding to inference, specifically by taking the content (not content sign) as a premiss and the 'message', i.e. the expression (i.e. linguistic sign), as the conclusion. This is at best an analogy, since genuine inference is from a set of sentences or thoughts to another sentence or thought. But even granted this analog, there is a problem: in what shape is the content possessed (as 'premiss') *before* it is encoded? Is it a piece of the world that is encoded? A piece of experience? Probably neither, since, prelinguistically, neither answers to verbalization. But the world or experience

as thought of is already relative to language. 'Encoding' is really speaking what is on one's mind. But the thought that is thus expressed is already in words albeit unspoken words. Peirce like Plato said we think in words. But how can our initial verbalization, whether overt or in thought only, be an inference *from anything*? Now Peirce did speak of our first judgments (perceptual judgments)—first in the sense of being inferred from preceding judgments—as being formed in a process that in some respects is *like* inference in that it is a limiting case of inference, but the type of inference he meant is abductive not deductive. The perceptual judgment or first verbalization is elicited by sensations, i.e. physiological stimuli, or by resultant sensory images, but it does not *encode* these: it is abductive precisely because it posits its own object, the supposed cause of the stimuli. This is crucial. As a first verbalization it does not put into words something not already in words. That cannot be done. Instead, first words introduce, create, posit their own objects. These may in some manner correspond to or fit antecedent realities, but as posited they are what answers to words. Hence, instead of deductive encoding we have abductive positing. And that, one could say, puts meaning where it belongs, viz. *in* language or as existing only as relative to language, and not in antecedent or prelinguistic reality or experience. Meaning, then, is inseparable from the interpretability of words. It involves potential interpretation. Reference is determined by the causal factors eliciting thought or utterances, and truth is a fit of meaning to reference, where 'fit' is to be explained in terms of the telos of speech.

Decoding, on the other hand, is interpretation of utterances by thought, which is *not* replacing expression by nonlinguistic content but is a translation of words into other words or, ultimately, it is a change in the interpreter's ('decoder's') emotions, actions, or dispositions to act. This is not to deny that the hearer must form a certain hypothesis about what the speaker meant (a genuinely abductive aspect of 'decoding'), nor the converse possibility that the speaker did not express himself fully or accurately.

This analysis would lead to the following postulates. There is no meaning apart from habits of interpretation. There are no linguistic signs that do not have a meaning. The supposed analogy between the expression system and the content system is really an analogy between the material aspect of language and its meaning, i.e. how its uses are to be interpreted. One could speak of the immediate object instead of the interpretant; but the dynamic object is the referent.

2. *Nominalism and realism in linguistics*

Philosophers have always thought of nominalism as a doctrine, not as a practice. They may therefore be excused for having trouble seeing the relation of nominalistic linguistics to the doctrine of nominalism, which is that the former is a way of doing linguistics to which doctrinal nominalists could not object, but that would seem deficient to those who are doctrinal realists. For if there are no classes in reality, but they exist in name only, as doctrinal nominalists claim, then any way of dividing up phenomena, including linguistic phenomena, is as good—or at least as true—as any other. And by 'nominalistic linguistics' I mean the practice of imposing an arbitrary taxonomy on linguistic phenomena.

This use of terms and concepts from the history of philosophy to make headway in linguistic theorizing may be interesting but also possibly confusing, the latter for the following reason. The linguistic phenomena classified might include linguistic universals (the Peircean 'types') as well as linguistic individual events (the Peircean 'tokens'). And one who is familiar with the nominalist/realist distinction as a matter of doctrine only might naturally suppose that by 'nominalist linguist' is meant one who denies the reality of linguistic universals. That, of course, would be an application of the nominalist *doctrine* to linguistic phenomena; but that, one can see now, is distinct from nominalist linguistics as a practice or method. Nominalism as a *practice* would not necessarily deny that universals are real; rather, it consists in deciding their classification arbitrarily—both their classification into subtypes, *if* they are segregated from individuals, and *whether* to so segregate them. Even their classification as real or unreal would be quite arbitrary.

The Chomskyan search for deep structure and generative principles looks relatively realist from a doctrinal point of view.[3] For whether or not surface phenomena are conceptualized in terms of types as well as tokens, the deep structure and principles look like universals, and especially so the way Chomsky and his followers speak of them. Chomsky and his school are nominalist linguists, not realist linguists, because their taxonomy of surface phenomena—the phenomena they wish to explain as following from deeper principles—is arbitrary. (It would follow that the hypothetical structure must be arbitrary too, for it is justified only by its capacity to explain those phenomena.)

[3]In using the label 'Chomskyan' I intend to let it refer not only to Chomsky himself (see now Chomsky 2000 for the latest tergiversations) but to all the latter-day offshoots of transformational-generative grammar as well—even those like Natural Phonology and Morphology or Optimality Theory (see Kager 1999 for a handy summary) that claim to be founded on principles that diverge from Chomskyan linguistics.

'Realism', of course, is used to designate the opposite of phenomenalism as well as the opposite of nominalism. With respect to doctrine exclusively, not method, Jakobson and his structuralist continuators (like Andersen and me) look like phenomenalists in contrast to Chomsky and his followers, since the former seem much more concerned with the description of what is here being called surface phenomena, whereas the latter plunge quickly to the (putative) underlying realities that explain them. One could say that Chomsky is in error for proceeding too quickly: after all, how can he abduce explanatory realities when he is wrong about the explanandum? But this is not so simple an issue as that. For if the classification of phenomena is to be real, not nominal, then it is often impossible to know what that classification is until the underlying realities have been identified. As an example from a domain other than language, consider whether it was possible to know that rusting, fire, and metabolism should be classed together as members of the same natural kind before they were all explained as different forms of oxidation. The circle here is like the hermeneutic circle: the explanans and the explanandum are found together, not first one and then the other.

But there is another way of looking at this which can be identified, mutatis mutandis, with that of semiotic neostructuralism in linguistics.[4] Realism in contradistinction to nominalism (doctrinally) is connected with teleology—or so, at least, Peirce appears to have thought. A natural class is one the members of which exist because each satisfies the same idea. That idea has a certain potency, and hence the class exists independently of anyone's having named it. This idea is consistent with the argument of the preceding paragraph according to which some natural classes may be those classes entailed by a true explanatory theory. But it is not limited to cases where the explanatory structures lie beneath the surface phenomena. Suppose language qua phenomenon has a history, and suppose that history can be understood by postulating goals not involving any underlying mechanisms. For example, linguistic change might be seen as tending toward a more adequate diagrammatization. Then we have a teleological basis for identifying natural linguistic classes, namely those that we have to attend to in order to understand language as diagrammatization. (This too involves a hermeneutic circle: neither the right description of the process nor the goal that explains it can be discovered without also discovering the other.)

If the preceding is a roughly correct account of the linguistic practice of semiotic neostructuralism, then it would seem that one who espouses the latter

[4]By 'semiotic neostructuralism' as applied to the study of language I mean the doctrine and method that emanate from an amalgamation of Jakobsonian linguistics with Peircean semiotics.

is in method, if not in doctrine, a realist as opposed to a nominalist, but a phenomenalist as opposed to a realist, and a teleologist.[5] One may doubt whether a semiotic neostructuralist is a phenomenalist in doctrine. For such a linguist does not deny, in fact, he presupposes that there are realities beyond or beneath language but for which his teleological account of linguistic change would make no sense. That is, there must be flesh-and-blood bodies that speak and listen, and it is their desires and needs that explain why ever more adequate diagrammatization is an inevitable if unintended goal. If the research program subtended by semiotic neostructuralism can be made to work, then it will indeed conflict with Chomskyan linguistics—and prove superior to it. Here is why.

Chomsky has a rather mechanistic view of language, for all that he understands that the freedom to compose sentences that are original, unpredictable, and yet intelligible is different from the unoriginal, predictable products of strictly mechanical action. His view is mechanistic nonetheless because he simply posits underlying structures by which sentences are to be generated. Possibly in a wider perspective, Chomsky is no more reductively mechanistic than a semiotic neostructuralist, in a wider perspective, is a phenomenalist. For he no doubt admits (or would admit) that the linguistic universals in our brains are not just there, period, but evolved, with the brain's evolution, as chance variants that were 'selected' by the principle of reproductive success. Similarly, the intentions or needs or felt urgencies to speak or to achieve certain outcomes might explain—but only in a context wider than Chomskyan linguistics—why language's generative mechanisms are used in this way rather than in that. But if we focus simply on the linguist's study, as diversely conceived by Chomsky and the semiotic neostructuralist, then there is this difference: for the one, the teleology of language is excluded from linguistic explanation, while for the other it is the very stuff of explanation. For the one, linguistic phenomena conform to a describable structure of highly abstract laws, while for the other linguistic phenomena exhibit an intelligible if less abstract, more complicated structure. For the one, the system is a given, and any changes in it are accidental, while for the other development is essential to language—development is more the reality than is any one system of rules—and that development is also intelligible and not merely given.

[5]T. L. Short (p.c.) points out that there *is* a methodological use of 'phenomenalist'; for instance, classical thermodynamics is often called 'phenomenal thermodynamics', not because its proponents are phenomenalists in philosophical doctrine but because it formulates the laws of thermodynamics without reference to the atomic theory of matter, which, with Boltzmann et al., was found to explain and quantify those laws.

That is the conflict. The reason the semiotic neostructuralist approach is, if it is successful, superior is that it can be used to explain the very evolution of the brain-mechanism or linguistic capacities and universals that Chomsky can at best describe. That is, given creatures somewhat sociable, exchanging signs as their way of life, then the survival value of their communicating more elaborate and precise diagrams would explain the retention of those fortuitous variations, say, in brain structure that promote exactly such powers of expressible diagrammatization. That is, the principle of this evolution will be itself linguistic, and continuous with the principles of postbiotic, strictly linguistic evolution. The thought here is not unlike that which refuses to postulate linguistic intentions separate from the capacity to exercise those intentions. Just as there could be no desire to speak without an ability to speak, so also there could be no evolution of linguistic capacities—even, or especially, at the physiological level—except among those who, already speaking to one another, will more likely survive as a species if they speak more effectively. Thus, instead of a neurophysiological explanation of language, we have a linguistic explanation of the higher cortex (and probably not just the speech centers either, since so many of our capacities for sensation and action would be bootless without our capacities for speech).

3. *Semiosis and linguistic change: efficient and final causation*

Peirce's distinction between legisigns and replicas can be used to good account in lifting some of the confusion that surrounds linguistic change, which is the end-directed evolution of a system of legisigns.[6] Replication is the end-directed use of already developed legisigns. In this process the legisigns (or rules of replica formation) do not function as efficient causes precisely: indeed, it is doubtful whether a rule or general type could ever be an efficient cause. But neither are they teloses (Gk. *telē*) of replication. The purpose of replication is communication (conveying information, issuing commands, expressing emotions, etc.). Thus, legisigns are not replicated simply for the sake of being replicated. They could be efficient causes of acts already explained by final causes—except for one thing. They could be efficient causes because final causes require the cooperation of efficient causes. Suppose I want Jones to close the door. I look around for means to do so. One means is replicating the English sentence, "Jones, close the door!" If that were the only means then, given my purpose, one can suppose that the availability of that legisign causes me (like a mechanical push) to replicate it. (But this is

[6]Conceptual change is the end-directed evolution of the rules of interpretation of symbols, sometimes with concomitant changes in the symbols themselves. Conceptual change then determines linguistic change, but in general this is not necessary to linguistic change.

wrong—why in a moment.) But the availability of alternative legisigns (e.g., "For God's sake, Jones, close the door!" or "Jones, dear fellow, I feel a draft.") means I must *choose*, and so those legisigns are not efficient causes. Legisigns cannot be efficient causes at all. In the first place, the efficient causes that must cooperate are those motor reflexes and the like that make my tongue wag, my mouth open and close, or my hand type these words. Secondly, legisigns are general types and hence can never be efficient causes. The upshot of this is that legisigns both exist for a purpose (they have evolved to make communication possible or to facilitate communication that was already possible) and are *used* when *we act* for the purpose of communicating. Thus, already existing legisigns are subsidiary final causes: we make such-and-so sounds or marks *in order to* replicate certain legisigns, and we replicate those legisigns *in order to* communicate something.[7] There is, therefore, an important difference between (1) legisigns developing and (2) legisigns being used.

Talk about final causation is often accompanied by contrasting references to efficient causation. An efficient cause is a particular event or condition that compels its effect. The effect follows the cause in accordance with a general law (a law of efficient causation). A final cause is not a particular event or condition and does not compel its effect. Suppose a man is seen bounding down a steep incline. Why? Possibly because the man was pushed. That would be an efficient cause. But perhaps the man acted in order to catch a goat. 'To catch a goat' is the final cause; it is not a particular event and did not compel the behavior.

Final causation is consistent with efficient causation, indeed requires it. Men cannot bound 'goatwards' if their muscles do not relax and contract, compelling movement of limbs. Presumably, then, the two types of cause explain different phenomena—or complementary aspects of the same phenomenon.

To explain something by a final cause is teleological explanation. Teleology is the doctrine that teleological explanations are sometimes legitimate, that some phenomena can only be explained teleologically, and that

[7]Notice that when we say things just for the sake of saying them, then legisigns may be truly final causes. But we need to distinguish three cases. The availability of certain meanings (= rules of interpretation of symbols) might intrigue me: so I want simply to express those ideas. Or it might be the legisigns themselves that intrigue me: poets (like the Russian futurist Mayakovsky) and composers (like Mozart) are said to have been fond of repeating certain (nonsense or foreign-language) words simply for the sake of their sound rather than their sense. Or it might be the truth we wish to state for *its* own sake, and in that case the final cause is the *agreement* of certain legisigns with an independent reality. In any case, replication of legisigns can be an end in itself, and in that case the legisigns are essential to one's ultimate purpose in speaking. That is to say, we would have a different purpose or none at all if we did not have those legisigns.

final causes are real. Teleological explanation was introduced deliberately by the Greek philosophers, primarily Plato and Aristotle, in explicit contrast to already well-established conceptions of causation—those that Aristotle identified as 'efficient' and that we can identify as 'mechanistic'. And already with Plato, it was recognized that this new form of explanation would be rejected by those who think (a) that everything can be explained by causes that compel or (b) that nothing that does not compel its effect could explain it.

In particular, what teleology was invented to explain is the existence of order—in human affairs, in individual actions, in plant and animal life, in the cosmos—wherever that order is inexplicable mechanistically. The point of teleology is to explain the emergence of order out of chaos. By contrast, the mechanistic world view of modern science admits none but efficient causes. However, not all forms of explanation in modern science conform to the mechanistic idea, even in its broadest and most up-to-date sense, but do approximate to the Aristotelian idea of explanation by final causes. Teleological theories are thus the best, or only, explanations of certain important classes of phenomena. Hence, we have good reason to suppose that final causes are real.

If this sounds too apodeictic for some readers' taste, it is probably due to the fact that teleology is badly understood.[8] An aid in dispelling some of the mist surrounding teleology is Peirce's idea of certain processes as 'finious', a neologism he coined for fear that "teleological is too strong a word to apply to them" (7.471).[9] These are nonmechanistic processes that "act in one determinate direction and tend asymptotically toward bringing about an ultimate state of things" (ibid.). The importance of nonteleological finious processes is that they explain how teleological phenomena are possible. One might say that they remove the mystery from teleology. Operating with the notion of finiousness imposes an obligation on the analyst: a hierarchical ordering of nonmechanistic explanations, some of which are merely finious, and some of which are teleological.

If one is to arrive at such an ordering following Peirce's conception, then it will be necessary to take into account his definition of final causation:

---

[8]Perhaps especially by linguists—like Lass (1997) and Labov (1994); see Short 1999 for a demolition of the former's antiteleological stance. As for the latter, his "Plan of the Work as a Whole", set out on the book's very first page, already betrays a fundamental misunderstanding of causation: it presents the organization of a projected three volumes into (respectively) "Internal factors", "Social factors", and "Cognitive factors"—as if these 'factors' were categorically distinct from each other (they are, of course, all 'internal').

[9]Citations in this form (volume and paragraph separated by a dot) are to Peirce's *Collected Papers*.

> ...we must understand by final causation that mode of bringing facts about according to which a general description of result is made to come about, quite irrespective of any compulsion for it to come about in this or that particular way; although the means may be adapted to the end. The general result may be brought about at one time in one way, and at another time in another way. Final causation does not determine in what particular way it is to be brought about, but only that the result shall have a certain general character. (1.211; cf. 1.204)

Any finious process is the result of fortuitous variation plus a principle of selection. These processes are everywhere observable in populations of individuals, whether molecules or living things. Other processes, equally finious, might be found within the actions of a single individual (not necessarily human).'[10] It is the nature of finious processes that their particular outcomes cannot be predicted; all that we can predict is their general tendency.

4. *Markedness in a theory of change*

Peirce understood a final cause as being a possibility—sometimes he said "idea", but that is not to be understood in a subjective sense as existing in some person's thought, that has a tendency to become actual, one way or another: "...every general idea has more or less power of working itself out into fact; some more so, some less so" (2.149).

It is in this sense that markedness must be viewed as a final cause in linguistic change.[11] When the question of causation is posed in terms of efficient and final causes—and teleological processes distinguished from finious—then the claim that, rather than markedness principles, it is "perceptual factors and processing strategies [that] may influence the development of linguistic structures" (Smith 2001:207) will be seen for what it is, a category mistake.

This mistake results from the apriorism that underlies how contemporary linguists commonly understand markedness (e.g., in Optimality Theory, but not only). On this view (partly reflected in Smith's contribution), markedness is simultaneously conflated with and pitted against notions like "sentence processing" or "perceptual strategies" as if markedness were an efficient cause, i.e. categorically of a piece with the latter. Lending support to skepticism regard-

---

[10] With respect to the deliberate conduct of human beings, the principle of selection is a type of outcome they have in mind, and which they consciously apply in choosing among the alternatives available to them. In other words, what we have in this case is purposefulness. Since an analysis of purpose would take us even farther afield, I refer the reader to the admirably clear exposé in Short 1999.

[11] In the event, I understand Andersen's conception of markedness (2001) to be compatible with this view. For a discussion of final and efficient causes in linguistic change that takes part-whole relations into account, see Shapiro 1991:16ff.

ing the relevance of markedness (and emanating directly from what I would now call the Apriorism Fallacy) is the perceived difficulty of assigning universal or immutable markedness values, even though markedness is invariably context-sensitive and dependent on the existence of choice between variants.

The question Why? as applied to linguistic change does not have a homogeneous answer. The problem of assigning markedness values is not solely the burden of linguists: it falls on language users as well. Linguistic data always contain the germ of ambiguity, of differing interpretations, and it is only by trial and error that the finious process of reaching a definitive markedness assignment proceeds. This process is necessarily always historical and not given apriori because at any given time linguistic habits, like all other habits, have a structure, and this structure is always *in statu nascendi*. But the important thing is that *an assignment will be reached*.

Language users do not need to wait for linguists to decide what is marked and what unmarked in order to be influenced by markedness considerations in making innovations and (tacitly) agreeing that some innovations qualify for the (social) status of full-fledged changes: they do it willy-nilly because they are impelled to by the power of the idea. Or as Peirce put it: "... it is the idea that will create its defenders and render them powerful" (1.217).

REFERENCES

Andersen, Henning. 1984. "Language structure and semiotic processes". *Arbejdspapirer, udsendt af Institut for Lingvistik, Københavns Universitet* 3.33–54.

Andersen, Henning. 2001. "Position paper: Markedness and the theory of change". In this volume, 21–57.

Chomsky, Noam. 2000. *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.

Jakobson, Roman & Morris Halle. 1971. *Fundamentals of Language*. Second edition. The Hague: Mouton.

Kager, René. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.

Labov, William. 1994. *Principles of Linguistic Change*. Volume 1: *Internal Factors*. Oxford: Blackwell.

Lass, Roger. 1997. *Historical Linguistics and Language Change*. Cambridge: Cambridge University Press.

Peirce, Charles Sanders. 1965–1966. *Collected Papers,* Volumes 1–8 ed. by C. Hartshorne et al. 2nd printing. Cambridge, Mass.: Harvard University Press.

Shapiro, Michael. 1991. *The Sense of Change: Language as History.* Bloomington: Indiana University Press.

Short, T. L. 1981a. "Peirce's concept of final causation". *Transactions of the Charles S. Peirce Society* 17.369–382.

Short, T. L. 1981b. "Semeiosis and intentionality". *Transactions of the Charles S. Peirce Society* 17.197-223.

Short, T. L. 1982. "Life among the legisigns". *Transactions of the Charles S. Peirce Society* 18.285–3 10.

Short, T. L. 1983. "Teleology in Nature". *American Philosophical Quarterly* 20.311–320.

Short, T. L. 1989, "Why we prefer Peirce to Saussure". *Semiotics 1988* ed. T. Prewitt et al., 124–130. Lanham, Md.: University Press of America.

Short, T. L. 1996. "Interpreting Peirce's interpretant: A response to Lalor, Liszka, and Meyers". *Transactions of the Charles S. Peirce Society* 32.488–541.

Short, T. L. 1998a. "Jakobson's problematic appropriation of Peirce". *The Peirce Seminar Papers* ed. Michael Shapiro, 3.89–123. New York & Oxford: Berghahn Books.

Short, T. L. 1998b. "What's the use?" *Semiotica* 122.1–68.

Short, T. L. 1999. "Teleology and linguistic change". *The Peirce Seminar Papers* ed. Michael Shapiro, 4.111–158. New York & Oxford: Berghahn Books.

Short, T. L. MS. "Peirce's theory of signs". To appear in *The Cambridge Companion to Peirce* ed. C. Misak. Cambridge: Cambridge University Press.

Smith, John Charles. 2001. "Markedness, functionality, and perseveration in the actualization of a morphosyntactic change". In this volume, 203-223.