

The Representation of Semiotic Structures in Language

Steven Bonta

Abstract

Based on previous published work on (Mandarin) Chinese, and following a discussion of the properties of the Peircean ontological Categories Firstness, Secondness, and Thirdness (as well as their “degenerate” or derivative versions) and their applicability to sign systems, in general, I examine evidence for paradigmatic and syntagmatic structuring, conditioned by these Categories, in Mandarin Chinese, Sora, Tamil, and Sanskrit, languages chosen because of the typological divergence amongst them, and because of the author’s familiarity with them and with their respective cultural milieus. The paradigmatic and syntagmatic structures identified arise in the presence of what I term positive and negative conditioning constraints arising from the Categories themselves, and which are shown to operate at three different levels in language, the morphosyntactic, the lexical, and the phonological. Because of this, a methodology grounded in Peircean semiotic structures is shown to have the explanatory potential to allow for a unified model of language structure, in general.

Keywords: Peirce, Peircean Categories, Peircean Sign, Peircean semiotics, semiotic structures, semiotic grammar, linguistic semiotics

1. Theoretical Underpinnings

1.1 The Peircean Categories

The basis of all cognition, i.e., of all intelligibility, must be some set of ultimate axioms, which approximate as nearly as possible the entelechy of Aristotle (“the

Universe qua fact”; Peirce, 1998, p. 304). The most fruitful such schema ever devised is the triad of primordial Categories elaborated by Peirce as the ontological scaffolding of what he styled the “phaneron”, that is, the continuum of cognizable reality. In a coming day, we hope, the Peircean architectonic will become an epistemological commonplace, like the Cartesian assumptions that have dominated most realms of academic discourse for the past several centuries. But for now, at least, Peirce’s most important insights remain somewhat at the periphery of most scientific and metaphysical avenues of inquiry, making the overview in this section a necessary service to the impartial reader.

Decades before Husserl and his phenomenology, Peirce had already concluded that the proper foundation for any architectonic or all-encompassing metaphysic was an examination of the most fundamental and general possible classes of being. While Peirce was far from the first metaphysician to attempt such an ambitious classification of ontological Ultimates, his system distinguishes itself for its breadth and simplicity. Peirce came to believe fairly early in his intellectual development that there were but three ultimate Categories of Being, and he termed them, prosaically enough, Firstness, Secondness, and Thirdness (see, e.g., “On a New List of Categories”, in Peirce, 1992, pp. 1-10, and “The Categories Defended”, in Peirce, 1998, pp. 160-178).

Put as concisely and generally as possible, a First or Firstness ([1]) is “the Idea of that which is such as it is regardless of anything else” (CP 5: 66). A Second[ness] ([2]) is “the Idea of that which is such as it is as being Second to some First, regardless of anything else, and in particular regardless of any Law, although it may conform to a law” (ibid.). A Third(ness) ([3]) is “the Idea of that which is such as it is as being a Third, or Medium, between a Second and its First” (ibid.). These three notions being absolutely and self-evidently primordial, it follows that all other phenomena—which we shall hereafter term “composite” or “non-primordial”—are subsumed by them, such that Firstness, Secondness, and Thirdness are often characterized in terms of some of their most familiar manifestations. For instance, Peirce was fond of likening Firstness to a “Quality of Feeling”, and pointing out that this Category encompassed, or was the dominant characteristic of, such non-primordial notions as spontaneity, freedom, variety, and freshness.

Following analogous lines of reasoning for the other two Categories, Peirce ascertained that Secondness embraces phenomena like reaction, resistance, otherness (Levinas’ “alterity”), compulsiveness, and corporeality, while Thirdness includes

mediation, habit, law, representation, plurality, evolution, continuity, and regularity among its many manifestations.

This taxonomy also gives rise to three “degenerate” Categories, a degenerate Secondness termed Firstness of Secondness, and two species of degenerate Thirdness, Secondness of Thirdness and Firstness of Thirdness. Each of these three degenerate Categories also embraces a range of phenomena, to be detailed hereafter.

It is Thirdness that confers upon reality its evolutionary character, as well as imparting its living, even volitive, substance. As a consequence, the full Categorical¹ accounting of reality is not of a mere congeries of Objects in blind reaction one with another, but of an infinite continuum of dynamic, purposive Events. And the essence of every Event is the Symbol, or, in the broadest sense, the Representamen, which we will here define as anything whose essential mode of being is representation, *sensu lato* (i.e., standing in a relationship between a First and a Second). We set aside here an obvious question that Peirce himself wavered on—namely: are all Representamens also Symbols?—except to observe that, in his later writings, he tended towards the affirmative, as his ringing declaration in one of his greatest papers attests.² At very minimum, Symbols are the best-understood of all Representamens. A natural proving-ground for Peirce’s system, therefore, is that grandest of all known symbolic systems, human language.

Much of Peirce’s output on the Categories was with reference to Signs, which he classified in alignment with the Categories themselves. Peirce’s most famous Sign taxonomy, the triad Icon—Index—Symbol, was grounded in the manner in which a Sign represents its Object. In the case of an Icon, a Sign represents its Object via some kind of qualitative similarity (a photograph, e.g.), i.e., via Firstness; an Index represents its Object by means of physical contiguity or some kind of deixis (a pointing finger, e.g.), i.e., a Secondness; and a Symbol represents its Object via some kind of habit, convention, or law, i.e., by virtue its being a Sign and nothing else (most words, e.g.)—a true Thirdness. Nor was this the only Sign taxonomy that Peirce produced. He also perceived that Signs could be classified according to their inherent nature, as Qualisigns (qualities or Firstnesses acting as Signs), Sinsigns (existents or Secondnesses as Signs), or Legisigns (laws or habits, i.e., Thirdnesses, as Signs). And Signs can be classified according to the way in which the cognizing Mind interprets the Sign’s relationship to its Object; if that relationship is interpreted as a First, then the Sign is a Rheme; if a Second, then the Sign is a Dicisign or Dicient Sign; and if a

Third, then the Sign is an Argument.

Not only that, the Sign itself, or, we may say, the semiotic Event corresponding to the Sign in its full sense, is also triadic, consisting of not only the Sign per se, but also its Object and its Interpretant, the last being a further Sign produced as an interpretive response to the Sign, usually a thought.

From Peirce's well-developed study of Signs, we can discern that the Categories were the ontological basis not only for their classification, but also for understanding their modes of operation. Because of the pervasive nature of Thirdness, for instance, the Sign can never be thought of, in any non-abstractive sense, as a discrete, self-contained object, but instead must be reckoned as an Event concatenated with an infinite succession of Events, cognizable but never dissociable from the infinite manifold of Sign-Events that constitute reality-qua-semiosis. This state of affairs suggests that triadic reasoning will yield a cognizable but non-atomistic description of any domain of reality.

Our aims in this paper are, firstly, to contribute to the further development of a methodical system of inquiry growing out of the Categories, and secondly, to apply it to aspects of human language, with a view to showing the explanatory power of semiotic structures. Indeed, a thoroughgoing methodology grounded in triadic reasoning is in order if Peircean realism is ever to meet the standards of modern scientific inquiry. The Categories themselves require a universe that is ordered, law-driven, and cognizable; yet the development of such a methodology was hampered by the fact that Peirce himself was (by his own admission; see "New Elements", Peirce, 1998, p. 309) not well grounded in linguistics, and died before the 20th-century revolutions in science and mathematics were truly underway. We suspect that, had Peirce lived to see the path-breaking linguistics of Sapir, Jakobson, Kurylowicz, and others; had been able to acquaint himself with mathematical innovations like tensors and manifolds; or had lived long enough to learn of the wonders of the quantum and relativistic realms of the physical universe, he would not only have imbibed deeply of these heady draughts, he would have greatly enlarged and refined his architectonic to accommodate such advances. Instead, it is left to those of us swimming in his vast wake to try to assemble piecemeal what he could have accomplished unaided. We note also that, pace Deely's reservations about a methodological approach to semiotics (Deely, 1990, pp. 11-13) should not be a deterrent, if our aim is to establish semiotics as a truly scientific enterprise.

Because triadic reality is constrained by Thirdness, it will always exhibit an element of regularity, i.e., it will be nomothetic, notwithstanding that this regularity ([3]) will always involve elements both of randomness (Firstness/[1]) as well as determinism (Secondness/[2]). The regularity being the basis for Thirdness and hence, for any interpretive Symbol (see, e.g., “Sundry Logical Conceptions”, Peirce, 1998, p. 369) implies that triadic reality may be represented as a systematic manifestation—yet the randomness begotten by Firstness also forces us to acknowledge that every triadic system will exhibit its irregularities. This, at least, is well-understood by linguists; Sapir’s famous dictum that “all grammars leak” becomes a guiding principle for all systemic representations of triadic semiotic reality.

1.2 Axioms

Before passing to the task of representing semiotic structures in languages, we must examine the Categories in some further depth, to infer from them as many axioms as we can concerning the characteristics of triadic reality. Since Firstness, Secondness, and Thirdness by definition encompass all phenomena, we can assemble from their essential properties a model of the metastructure of reality, or any subdivision thereof, with the aim of informing the development of our aforementioned methodology. First of all, reality must be triadic, as Peirce so often insisted; this has been so thoroughly and persuasively demonstrated by Peirce that no further elaboration is required. A corollary to the necessity of the Triad is, that reality must be composite, in some sense yet to be made precise.

Secondly, all reality must be hierarchical, or in other words, the general notion of priority (or, in Peirce’s terminology, “gradation”; see Peirce, “On a New List of Categories”, Peirce, 1998, p. 1) must obtain wherever Peircean “phaneroscopy” may cast its beam of inquiry. That this must be so can be readily appreciated, not only from the immanent priority of the primordial Categorical triad itself, *vis-à-vis* all other possible categories of composite phenomena, but also from the relationships of the Categories among themselves, namely, the fact that the Categories themselves are unequal, as implied by their enumerative designations. A First/[1] is, in some sense, prior to a Second/[2], whether in the sensual or in the conceptual realm, and a Third/[3] is subsequent to both [1] and [2]. At the same time, inasmuch as [2] always involves [1], but not the reverse, and [3] always involves both [1] and [2], but not the reverse, [2] is superior, *sensu lato*, to [1], and [3] to both [2] and [1].

Thirdly, all existent and nomothetic reality (i.e., reality of which [2] and [3] is constitutive, respectively) is inherently non-commutative. This follows from the preceding axiom, and also from the facts that a) every [2] must involve a [1], but not the reverse, and b) every [3] must involve [2] and [1], but not the reverse. That is:

$$\begin{array}{ll} [1] \in \{[2]\} & [2] \notin \{[1]\} \\ [2], [1] \in \{[3]\} & [3] \notin \{[2]\}, \{[1]\} \end{array}$$

Here we have taken slight liberties with mathematical notation, for want of better alternatives; it will be appreciated that, e.g., $\{[2]\}$ is to be understood to mean “the set of all phenomena constitutive of [2]/Secondness”.

The foregoing discussion makes use of Peirce’s term “involve”, which carries a connotation of interiority, also *sensu lato*.; that is, to say [2] “involves” a [1] is as much as to say that there can be no alterity without identity, no Index without Icon, no Other without Self, etc., and that, in each instance, [1] is, in one way or another, constitutive of [2]. Similar arguments can be made for [3] with respect to both [1] and [2]. We shall say that [1] *inheres in*, or *is inherent in*, [2], and that both [1] and [2] *inhere in*, or *are inherent in*, [3], but not the reverse. As for each Category taken absolutely in and of itself, representing, *in abstracto* at least, its own Universe of Being, we may refer to its *immanence* or *immanent characteristics*.

But now arises a question that needs clarification: is there any non-constitutive implicational sense by which [1] requires [2], or [2] requires [3]? To answer this, we must try to think, first, of the characteristics of reality borne out by the Categories. From [1], we infer that reality is stochastic, or, that there is an element of chance in reality. This ontological posture Peirce termed tychism, and the doctrine of tychism also begets such abovementioned qualities as variety and freshness. From [2], we infer that reality is deictic, requiring at minimum a dyad. This notion encompasses such fundamental phenomena as duality, determinism, discreteness, and compulsiveness. From [3], we infer that reality is representational, meaning that it serves to relate a [2] with a [1]. Representation/[3] gives rise, in turn, to the notions of plurality, thought, infinity, continuity, habit, and law.

We now consider the nature of [1], taken by itself. Given that both [2] and [3], which are constitutive of all real phenomena, both require [1] as an inherent, it follows that the only conceivable context for [1] dissociated with [2] and [3] is

abstractive—but every abstraction requires an act of cognition, which is perforce a [3]; and so every [1] must also require [2] and [3], but not in a constitutive sense. Instead, this requirement is both relational (a Third) and exterior (a Second). To characterize this relationship, we say that [3] *adheres to*, or *is adherent to*, [1], and we denote it as follows:

$$[3] \alpha [1]$$

Because all acts of deliberative abstractive cognition also require an Index (a Second), we may also assert that

$$[2] \alpha [1]$$

and

$$[3] \alpha [2].$$

Only in this way can we understand how it is that, e.g., a First can have representational force, as an Icon, Qualisign, or Rheme; the representational element, in such instances, is adherent to the sign.

This line of reasoning gives rise to another fundamental property of reality, namely, that representational phenomena may be immanent, adherent, or inherent and that, inasmuch as the abstractive act described previously may apply to any of the Categories, the triad immanence/inherence/adherence may apply to any element of reality. As a consequence, any sign may be regarded either as an independent thing in and of itself/[1], as standing in relation to some *ens exterior*/[2], or as a composite of inherents/[3]. From this we conclude that all reality is implicational, whether via immanence, adherence, or inherence.

From the composite, hierarchical, and non-commutative characteristics of reality we infer the universality of the *syntagm*, that is, the ordered sequencing of all constituents of the dynamic reality constitutive of all events. In linguistic semiosis, the syntagm is manifest as morphosyntax. All syntagms are representational.

Finally, we may infer from all of the foregoing that all reality is classificatory. This follows from the originary character of triadic reality *qua* classification, and is given additional vigor by such notions as plurality and law (Thirds). Any classification

being itself a sign, it follows that it may be predicated either upon immanence, adherence or inherence, or any proportion of these in combination with the others. In practice, of course, human classification, via language, art, or culture in general, always involves some degree of adherence, or, as we have elsewhere styled it, lensing, whereby a particular Category-sign is given representational primacy (Bonta 2018). In such instances, then, [1] may be selected to represent [2], [23], [3], etc. More concretely, we might find, for example, that whereas “music,” taken with respect to other members of the set of “things constitutive of a culture-sign” (which set might also include science, mathematics, engineering, religion, philosophy, clothing, architecture, language, etc.), would be regarded inherently as a First (where, e.g., sculpture might be a Second and mathematics a Third), yet music might represent, by adherence, a Second or a Third, in addition to, or rather than, a First. After all, music may be evocative of a pure quality of feeling ([1]), but it may also be evocative of sensual or compulsive action ([2]), or even abstractive cognition ([3]).

From the classificatory character of triadic reality we are led ineluctably to the notion of the *paradigm*, the ordering *in potentia* of all elements of a related class, among which obtain relationships of both similarity and priority.

Leaving aside the question of the primordially and universality of representation, we assume, based on the foregoing arguments, that semiosis—like all other phenomena grounded in the Categories—represents itself structurally (both syntagmatically and paradigmatically), and that the task of classifying and evaluating semiotic structures is best undertaken with the Categories themselves as the basis thereof. Any methodology of classification perforce involves a system, and, in this case, any such system should be an Index involving an Icon, writ large, suggestive of the manner in which the Categories themselves have been shown to operate.

The principle of priority described previously is made explicit in the linguistic notion of a mark, or markedness; for, just as a primordial hierarchy is implied by the relationship of priority among the constituents of the primordial Triad, so too we expect to find groups of semiotic objects of inquiry that may be differentiated hierarchically, based on the presence or absence of a mark or marks, whose presence we shall denote by “+”. Since we are using the Categories themselves as our basis, we assume that they themselves may constitute the marks referred to, as, e.g., [+1], [+12], etc., or in other words, that any conceivable mark, whatever its explicit contours, may—like all non-primordial phenomena—be represented in terms of the Categories.

That this is admissible is easily demonstrated. Assume we have a mark of any type, except that it cannot be a Category, pure or degenerate, per se, and let us denote it $[+\mu]$. Assume further two Objects, X and Y, brought into a classificatory relationship or set, and that one of them, X, does not involve $[+\mu]$, while the other, Y does. This relationship we may denote as $\{X, Y[+\mu]\}$. It will be appreciated that, no matter the immanent character of $[+\mu]$, within this classificatory set it operates as some Quality, i.e., a Firstness, of Y with respect to X. That is, [1] is immanent in $[+\mu]$, and as such, it may be represented as [+1]. But as we have already seen, every act of classification involves representation, which for our purposes requires some act of cognition. This being the case, both [3] and [2] are adherent to $[+\mu]$, and could therefore be likewise represented as its embodiment, every act of cognition being of the nature of a Symbol, and the intrinsic nature of a Symbol being pure representation without reference to any iconic or indexical constraint. Put simply, anything of the nature of a Symbol represents purely by being so regarded, and for no other reason (“New Elements”, Peirce, 1998, p, 321).

Now a mark having the nature of a Symbol, albeit a rudimentary one, it follows that it will also involve an Index, which, by being deictic, will have the effect of limiting or establishing boundary conditions for the class of Objects thereby denoted, with respect to some other, more general Symbol. Consider the Word-Symbol “mammal”, denoting a certain type of creature within the larger class denoted by ‘animal.’ That ‘mammal’ is marked with respect to “animal” may be seen by the Index embracing the set of characteristics that differentiate mammals from all other members of the animal kingdom. This composite Index serves to call attention to those differentiating characteristics (which, in the manner of all triadic semiosis, are neither absolute nor exceptionless, as with the case of outliers like the bat and the platypus).

But as we have seen, the Categories themselves, including degenerate forms, may serve as marks. This means that any of them may be enlisted, by adherence, to serve as an Index determining the boundary conditions for some class of non-primordial phenomena qua Objects. In other words, an act of cognition (a Third) may assign, e.g., [+12] (i.e., the degenerate Category Firstness of Secondness) as a mark (a Second denoting a First) to some congeries of Objects, thus transforming them into a class, i.e., a single composite Sign. But because Thirds embrace both the notions of habit and generality, it follows that any such Third may spread from one Mind-Symbol to the

next (creating a community of Mind-Symbols wherein this particular Symbol-as-Mark becomes established habit), and may be diffused outward from its original domain of application to approach the entire continuum of cognizable Being (see “The Law of Mind”, Peirce, 1992, pp. 312-333, for Peirce’s famous description of this process). In this way do Culture-Signs come into being; the process by which one or more Category-Symbols comes to serve as a maximally general signifying Index by which all reality is represented is called lensing, as already mentioned. In the foregoing example, we say that the lens is [12].

Writ large, lensing is the way in which a Mind-Symbol “reduce[s] the manifold of sensuous impressions to unity”, per Peirce (“On a New List of Categories”, Peirce, 1992, p. 1). It is a universal condition imposed upon the mind by what we might style the “tyranny of Secondness”, the all-encompassing constraint operative in a corporeal Universe. It is Secondness, more than the other Categories, that foremost urges itself upon the senses. This tyranny is responsible for the widespread fallacies of materialism and kindred schools of thought that cannot see beyond tangible reality.

This being the case, it follows that our primary objects of inquiry in any semiotic system must be *explicit*, or in other words, must be elements with Secondness strongly inherent. The explicit aspects of language include not only the morphosyntactic “surface structure”, but also the lexicon and phonology. These all manifest explicit semiotic structures that have been selected by a given speaker community to represent the “manifold of sensuous impressions”. Morphologizing is a well-studied process by which Word-Symbols (or lexemes) are selected as Indexes signposting morphosyntactic semiotic structures. This line of reasoning gives rise to the notion of the *morphosemiotic*, that is, that all form is associated, however haphazardly, with meaning, and that syntactic, morphological, lexical, and phonological structures are, at root, semiotic in nature—indeed, “structure” as such is the Indexical manifestation of semiosis. Put otherwise, the assumption of the morphosemiotic implies that explicit morphology confers semiotic prominence, i.e., signifies semiotic priority.

Our hypothesis is that study of these morphosemiotic structures in any language will reveal the Categorical lens associated with that language, and that, inasmuch as that lens operates after the manner of a Symbol, it will impose structural unity and consistency across the entire composite (but withal interconnected) Language-Sign (i.e., language qua sign).³ Further, because every semiotic lens is associated with one or more Categorical marks⁴, we propose that in any Language-Sign, these Categorical

marks act as constraints, in the same manner as other linguistic constraints, but at a maximally general, primordial level. This means that we may have both positive and negative constraints; returning to the foregoing example, [+12] is a positive constraint—but it also presupposes [-3], a negative constraint. These constraints are expected to condition both syntagmatic and paradigmatic structures. In the following case studies, we will see more concretely how both positive and negative Categorical constraints are manifest, in both paradigmatic and syntagmatic contexts, at multiple levels of language.

2. Case Studies

2.1 Mandarin Chinese and [12]

The reader will already have appreciated the potential complexities and pitfalls of semiotic analysis along the lines we are proposing. The very nature of triadic reality—evolutionary, continuous, and often stochastic—tends to generate exception-riddled complexity whereof the regularity can be very difficult to discern. Signs, including Category-Signs, often blend and hybridize, making it extremely difficult to tease out the semiotic properties of each constituent Sign from the amalgam. We seek to reduce the complexity of the subject matter while ensuring, as far as possible, that our own observations are accurate. The only way to do this is to seek languages that have proven less susceptible to foreign influence, and to limit our inquiry to languages that we have some familiarity with. Otherwise put, relying on data from unfamiliar languages or from languages with strong borrowed elements (like modern English) is not the best way to establish the theory. Instead, I will rely only on languages and language areas where I have significant personal expertise, from regions that have proven to be relatively opaque to foreign importations. Our approach will be to examine each language at the phonological, lexical, and morphosyntactic levels, and identify the dominant typological characteristics at each level. We will examine those characteristics in light of the Peircean Categories, to see whether a particular Category (pure or degenerate) appears to predominate. As we have already tested this method in previous work with Mandarin Chinese (Bonta, 2020), we summarize the results of that study first, to show how this method may yield fruit. In subsequent sections, we will apply it, *mutatis mutandis*, to other languages.

In previous work, I found Mandarin Chinese (hereafter “Chinese”) to be subject

to [+12], or Firstness of Secondness, as the primary constraint. To see how this has proven to be the case, we first must examine the characteristics of [12], one of Peirce's three "degenerate" Categories. In the case of [12], as also the other two degenerate Categories [13]/Firstness of Thirdness and [23]/Secondness of Thirdness, Peirce—in the portions of his writings that have so far been published—is comparatively reticent, reserving most of his exemplifications and in-depth descriptions for the "pure" Categories Firstness, Secondness, and Thirdness. The result is that, whereas we can rely comfortably on Peirce's characterizations of the latter three, the former three must, in many respects, be fleshed out by modern investigation. This, it seems, is a major uncompleted task of Peirceana inasmuch as, from the linguistic evidence to be presented herein, the degenerate Categories appear to be at least as prominent in semiotic structuring as the pure Categories, and perhaps more so. For example, based on our results from Chinese, we would anticipate that [12] is a prominent (though perhaps not exclusive) structuring constraint in many of the major languages of southeast Asia, of the Tibetan plateau, and of China itself, since these languages all bear significant typological similarities with Chinese in some or all of the criteria that will be shown following.

Peirce characterized Firstness of Secondness/[12] as follows:

Category the Second has a Degenerate Form, in which there is a Secondness indeed, but a weak or Secondary Secondness that is not in the pair in its own quantity, but belongs to it only in a certain respect. Moreover, this degeneracy need not be absolute but may be only approximative. Thus a genus characterized by Reaction will by the determination of its essential character split into two species, one a species where the Secondness is strong, the other a species where the Secondness is weak, and the strong species will subdivide into two that will be similarly related, without any corresponding subdivision of the species. For example, Psychological Reaction splits into Willing, where Secondness is strong, and Sensation, where it is weak; and Willing again subdivides into Active Willing and Inhibitive Willing, to which last dichotomy nothing in Sensation corresponds. (CP 5: 69)

In another paper, Peirce gave a few more details about the nature of Firstness of Secondness with respect to "pure" Secondness:

[T]here is a degenerate sort [of Secondness] which does not exist as such, but is only so

conceived. The medieval logicians (following a hint of Aristotle) distinguished between real relations and relations of reason. A real relation subsists in virtue of a fact which would be totally impossible were either of the related objects destroyed; while a relation of reason subsists in virtue of two facts, one only of which would disappear on the annihilation of either of the relates.... Rumford and Franklin resembled each other by virtue of being both Americans; but either would have been just as much an American if the other had never lived. On the other hand, the fact that Cain killed Abel cannot be stated as a mere aggregate of two facts, one concerning Cain and the other concerning Abel. Resemblances are not the only relations of reason, though they have that character in an eminent degree. Contrasts and comparisons are of the same sort. Resemblance is an identity of characters; and this is the same as to say that the mind gathers the resembling ideas together into one conception. Other relations of reason arise from ideas being connected by the mind in other ways; they consist in the relation between two parts of one complex concept, or, as we may say, in the relation of a complex concept to itself, in respect to two of its parts.... But [all these relations] are alike in this, that they arise from the mind setting one part of a notion into relation to another. All degenerate seconds may be conveniently termed internal, in contrast to external seconds, which are constituted by external fact, and are true actions of one thing upon another. (CP 1: 365)

In my own recent work, I offered the following clarification:

Robertson (1994), in his examination of English verb affixal morphology in light of the Peircean Categories, characterizes Firstness of Secondness as existing “where the mind sees a single notion as two, or where one of the two objects is vague and the other a more focused version of the other”.

In other words, Firstness of Secondness always involves an unequal dichotomy. It is the result of the mind resolving a single idea into two objects, complementary or opposing. We might add that, just as a pure Firstness must be understood to absolutely exclude all Seconds and Thirds (CP 1: 358), a Firstness of Secondness will, by its very nature, tend to exclude Thirdness altogether. (Bonta, 2020)

Our task, then, was to examine the lexicon, the morphosyntax, and the phonology of Chinese, to see whether a) the notion of the “unequal dichotomy” characteristic of [12] is given any prominence and b) whether Thirdness/[3] is in any discernible way

minimized. In other words, we looked for lexical, morphosyntactic, and phonological evidence for the constraints [+12] and [-3].

2.1.1 *The Chinese lexicon*

Consider first the Chinese lexicon as a starting point for looking for conditioning constraints. There are three general conditioning features of Chinese lexical entries, all of them conspicuous and readily noticed by any student of the language. These are: first, the relatively small number of admissible syllables in Chinese (no more than approximately 1300 in all, or 400 if tonal distinctions are ignored—this in comparison to at least 10,000 in English); second, the fact that every syllable has full lexemic value (leading to an enormous number of homophones in Chinese); and third, the fact that the overwhelming majority of actual Chinese words in common usage are dilexic, that is, they consist of two syllables, each of which is represented by one written sign and which corresponds in its own right to an independent lexeme.

Considering first the paucity of admissible syllables, and the fact that every such syllable represents a full lexeme, we observe that the organization of sounds into syllables and words, in general, tends to beget enormous variety, at least in English. This is because the relationship between sound and meaning is primarily Symbolic, or in other words, premised upon a given sound configuration's being assigned a meaning and upon nothing else. It is the immanently arbitrary character of Symbols, in general, that tends to impart diversity, both in length and in phonological complexity, to a lexicon. In English, for example, if we take a fairly random sampling of well-known mammals, we may note significant differences in sound shape, syllable count, and the like: {*elephant, tiger, giraffe, rhinoceros, moose, antelope, hippopotamus, lion, bear, wolf*}. None of these terms is analyzable in any way (aside from foreign roots, e.g., the Greek roots evident in *rhinoceros* and *hippopotamus*), i.e., the 'gi-' of 'giraffe', the 'ti-' of 'tiger,' the '-phant' of 'elephant', etc., have no independent significance; they cannot be separated from the rest of the respective Word-Symbols to which they belong.

But the situation in Chinese is entirely different. The equivalent Chinese set for the above animal words is as follows: {*dà xiàng, lǎo hǔ, cháng jīng lù, xī niú, tuó lù, líng yáng, hé mǎ, shī zi, xióng, láng*}. Every one of the syllables in the preceding set is meaningful, with the literal meanings for the multisyllabic entries being as follows:

<i>dà xiàng</i>	<i>'big elephant' ('elephant')</i>
<i>lǎo hǔ</i>	<i>'old tiger' ('tiger')</i>
<i>cháng jǐng lù</i>	<i>'long neck deer' ('giraffe')</i>
<i>xī niú</i>	<i>'rhinoceros cow' ('rhinoceros')</i>
<i>tuó lù</i>	<i>'camel deer' ('moose')</i>
<i>líng yáng</i>	<i>'antelope goat' ('antelope')</i>
<i>hé mǎ</i>	<i>'river horse' ('hippotamus')</i>
<i>shī zi</i>	<i>'lion offspring' ('lion')</i>

Leaving aside for a moment the obvious superfluities of some of these words (e.g., Why ‘old tiger’, if the lexeme *hǔ* means ‘tiger’ by itself?), we see that, in stark contrast to English, Chinese has no undecomposable polysyllabic words. Note also that nearly all of the above monosyllabic lexemes have multiple possible meanings, typically indicated by a range of different written signs. For example, the lexeme-syllable *shī*, ‘lion’ (represented by the sign 狮) may also signify ‘err, make mistakes’ (失), ‘master, teacher’ (师), ‘poem, poetry’ (诗), ‘carry out, execute’ (施), ‘damp, wet’ (湿), and ‘corpse’ (尸). All of these lexemes are identical as to their sound shape (including the tone; if we allow for the many additional possibilities where tone alone is differentiated, we will arrive at dozens of homophones or near-homophones). Only in writing may these words be differentiated, via the different form of the characters used to represent them. And we would obtain similar results for every single lexeme-syllable in the preceding list of animal names, and, indeed, for each of the roughly 1300 allowable syllables in the Chinese syllabic inventory.

The effect of this peculiar lexemosyllabic typology is to reduce, as far as possible, the allowable number of sound configurations participating in word formation. Limiting as far as possible—while still maintaining intelligibility—the number of available syllables, in combination with disallowing any unanalyzable polysyllabic lexemes, reduces as far as possible the arbitrary relationship between sound and meaning. Otherwise put, reduction of words in Chinese to analyzable combinations of a comparatively small set of syllables has diminished as far as possible the arbitrary—and therefore the symbolic—aspect of word formation (without, of course, obliterating it entirely, since linguistic signification without some degree of symbolism is impossible). All Symbols embodying Thirdness/[3], as we have seen previously, may represent this minimizing of pure Symbolism as the negative constraint [-3]

pervasively conditioning the Chinese lexicon.

As for the positive constraint [+12], the defining typological characteristic of the Chinese lexicon, as the reader may have noticed from the aforementioned set of animal names, is its overwhelming preference for disyllabic (and therefore dilexic) words. This preference embraces not only nouns, verbs, adjectives, and adverbs, but also extends even to conjunctions and prepositions. In every word class, monosyllabic/monolexic entries may also be deployed, but they tend to be far less commonly used than dilexemes, and far less frequent in informal speech than in more literary contexts. This is in part because the preponderance of dilexemes in the Chinese lexicon is very much an active evolutionary process, or, we might say, a very conspicuous operation of the Symbolic Third bringing Chinese into ever-more exact conformity with its conditioning lens.

The phenomenon of alternative monosyllabic and disyllabic words in Chinese is well-known, and is referred to as “elastic word length”. Duanmu (2013) found that 80 to 90 percent of all Chinese words have elastic length, with 92 percent of nouns and 83 percent of verbs in the sample he worked with displaying this characteristic.

Dilexic words may conform to any of a number of common patterns, including figurative oppositions (e.g., *dōng-xī*, ‘thing’, literally, ‘east-west’, or *shēn-qǐǎn*, ‘depth’, literally, ‘deep-shallow’), redundant pairings (e.g., *bāng-zhù*, ‘help’, literally, ‘help-help’ and *yán-jiū*, ‘research’, literally, ‘research-research’), phrasal verbs that incorporate a direct object (e.g., *zuò-fàn*, ‘cook, prepare food’, literally, ‘make-food’), and pairs involving a “light” syllable with no literal meaning that serves to do nothing more than create a word in its “full form” (e.g., *dāo-zi*, ‘knife’, whereof the second element, *-zi*, simply means ‘offspring’ in its original, literal form, and is now used to create hundreds of full-form Chinese words out of original lexemes—in this case, *dāo*, ‘knife’—which were the historical words in and of themselves, and which still may be found as monosyllabic lexemes in very formal or literary contexts). Other examples of “light” syllables include *-shì*, ‘be, is, are’, *lǎo-*, ‘old; venerable’, and *dà-*, ‘big’. A few examples of disyllabic words using these light syllables are shown on Table 1:

Table 1. Disyllabic words with “light” syllables *-zi*, *-shì*, *lǎo-*, and *dà-*
lǎo-hǔ, ‘tiger’ (*hǔ*, ‘tiger’)
lǎo-yīng, ‘eagle’ (*yīng*, ‘eagle’)
lǎo-wài, ‘foreigner’ (*wài*, ‘outside[r]’)

hóu-zi, ‘monkey’ (*hóu*, ‘monkey’)
shì-zi, ‘persimmon’ (*shì*, ‘persimmon’)
yàn-zi, ‘swallow’ (*yàn*, ‘swallow’)
zhú-zi, ‘bamboo’ (*zhú*, ‘bamboo’)
dà-xiàng, ‘elephant’ (*xiàng*, ‘elephant’)
dà-gǒu, [colloquial] ‘dog’ (*gǒu*, ‘dog’)
dà-nǎo, ‘brain’ (*nǎo*, ‘brain’)
dàn-shi, ‘but’ (*dàn*, ‘but’)
hái-shi, ‘still’ (*hái*, ‘still’)

Some examples of disyllabic words involving redundant pairings of lexemes are shown, along with both translations and literal glosses, on Table 2:

Table 2. Disyllabic words with redundant pairs of lexemes

<i>yán-jiū</i>	‘research-research’ (research)
<i>wú-gōng</i>	‘centipede-centipede’ (centipede)
<i>bāng-zhù</i>	‘help-help’ (help)
<i>huáng-dì</i>	‘emperor-emperor’ (emperor)
<i>pī-píng</i>	‘criticize-criticize’ (criticize)
<i>tiān-kōng</i>	‘sky-sky’ (sky)
<i>lián-hé</i>	‘unite-unite’ (unite)
<i>wū-yā</i>	‘crow-crow’ (crow)

This love of dilexemic, disyllabic words in Chinese has arisen as a result of a compulsive need to create complementary pairings at the lexical level as often as possible. But whence arises this need in the first instance? In previous work (Bonta, 2020), we argued that this conspicuous and pervasive typological characteristic of the Chinese lexicon is a consequence of the positive conditioning constraint [+12], that is, the lens of Firstness of Secondness. Recall that Firstness of Secondness is characterized by “an unequal dichotomy... the result of the mind resolving a single idea into two objects, complementary or opposing”. The dichotomous character of the overwhelming majority of Chinese lexical entries can scarcely be questioned on the evidence; but in what sense may the opposing or complementary entries of a given dilexemic Chinese word be interpreted as unequal? In the first place, owing to the immanent nature of the Categories already set forth (see Peirce, “On a New List

of Categories”, Section 2), we would expect, for *any* taxonomy arising directly from a Categorical lens, a self-evident hierarchy or gradation. The “unequal dichotomy” is simply this hierarchical character manifest in [12]. In the case of dilexic Chinese words, we can easily observe abundant evidence for the “unequal dichotomy” as a manifestation of [+12]. For example, a large number of dilexic Chinese words drop the tone on the second entry, suggesting iconically a privileging of the first over the second syllable/lexeme. Some examples include pairs in *-shì*, like *dàn-shì* and *hái-shì* noted previously, in which the tone of *-shì* is neutralized, as well as many other such pairings (in general, the more common the word, the more likely such tonal neutralization will take place).

Perhaps even more tellingly, Chinese frequently forms new disyllabic compounds out of existing disyllabic words, by dropping one of the two syllables in each entry. The resulting compound then furnishes strong lexical evidence that the syllable/lexeme from each original word chosen to participate in the resulting compound is, in fact, the semantically “dominant” or prior entry of the original word. And this notion is given additional credence by the fact that, in cases where the same word forms a multitude of different compounds, it is usually the same syllable that is chosen to represent the original disyllabic word in each of the resultant compounds. Table 3 furnishes several examples to illustrate how this process works:

Table 3. Chinese compounds formed from reduced pairs of dilexic words

Word	Meaning	Full forms (+ meanings of original words)
<i>dì-tiě</i>	‘metro’	<i>dì-xià – tiě-lù</i> (<i>dì-xià</i> , ‘underground’; <i>tiě-lù</i> , ‘railroad’)
<i>gāo-tiě</i>	‘bullet train’	<i>gāo-sù – tiě-lù</i> (<i>gāo-sù</i> , ‘high-speed’; <i>tiě-lù</i> , ‘railroad’)
<i>huá-shè</i>	‘Chinese community’	<i>huá-rén – shè-qū</i> (<i>huá-rén</i> , ‘Chinese’; <i>shè-qū</i> , ‘community’)
<i>hé-wǔ</i>	‘nuclear weapon’	<i>hé-zi – wǔ-qì</i> (<i>hé-zi</i> , ‘nucleus’; <i>wǔ-qì</i> , ‘weapon’)
<i>huán-dài</i>	‘pay off a loan’	<i>cháng-huán – dài-kuǎn</i> (<i>cháng-huán</i> , ‘repay’; <i>dài-kuǎn</i> , ‘loan’)

It should now be evident from the evidence given that the Chinese lexicon is pervasively and consistently conditioned both by the positive constraint [+12] and the negative constraint [-3]. Space will not allow a fuller exploration of the lexical evidence for these two constraints, but much more evidence might be adduced in a

fuller treatment of this phenomenon in future research.⁵

2.1.2 Chinese morphosyntax

The fundamental purpose of language being the construction of communicative propositions, it follows that the most essential characteristic of any language is the way by which it unifies the two fundamental parts of any proposition, the Subject and the Predicate. Regarding the semiotic sense of these two terms familiar to every grammarian, Peirce pointed out that the Subject, writ large, consists of the set of Indices proper to a given predicate, which in turn encompasses the entirety of ideas involved in the verb or other grammatical predicate:

[I]n order properly to exhibit the relation between premises and conclusion ... it is necessary to recognize that in most cases the *subject index* is compound, and consists of a *set* of indices. Thus, in the proposition “A sells B to C for the price D”, A, B, C, D form a set of four indices. The symbol “___ sells ___ to ___ for the price ___” refers to a mental icon, or idea, of the act of sale, and declares that this image represents the set A, B, C, D, considered as attached to that icon, A as seller, C as buyer, B as object sold, and D as price. If we call A, B, C, D four *subjects* of the proposition and “___ sells ___ to ___ for the price ___” as predicate, we represent the logical relation well enough, but we abandon the Aryan syntax. (“Of Reasoning in General”, Peirce, 1998, pp. 20-21)

This being the case, there are, broadly speaking, three morphosyntactic strategies deployed by languages to effect the colligation of Subject and Predicate, as we have noted previously (see Bonta, 2018): agreement, contiguity, and blending. Any of these three strategies may represent the Symbol of assertion, whereby the mind brings into relationship Subject and Predicate, by affirming, in effect, that the set of Indices corresponding to the semiotic Subject are, indeed, related to the Icon or mental image associated with the Predicate. The Symbol of assertion is above all a mental act, or Symbol of consciousness; it is often, although not always, signaled overtly by a copula or by some sort of verb inflection. Note that the foregoing is a semiotic typology, not a strictly formal grammatical one, and as such will not always neatly accommodate the various morphosyntactic typologies, like isolating and agglutinative, favored by linguists.

Blending refers to the formal conflation of morphemes representing Subject

and Predicate into a single lexeme, and is especially conspicuous in so-called polysynthesis (the creation of extremely long and complex “sentence words”, widely attested in Native American and Australian Aboriginal languages, among others), but is also present to some degree in agglutinative and fusional languages. Blending, as we have elsewhere shown (Bonta, 2018), is meaning association qua Firstness, because it seeks to merge the pair Subject and Predicate into an undifferentiated singularity ([+1]), and also because it minimizes lexical boundaries, contrasts, and distinctions ([-2]).

Contiguity refers to the strategy of juxtaposing Subject and Predicate temporospatially, such that the mind is directed to associate the two; this process is tantamount to meaning association via Secondness, because it relies critically on deixis for its effect. In so-called analytic or isolating languages, this strategy is especially conspicuous.

Agreement, as we have elsewhere noted, “involves the purely symbolic relationship between word pairs, like subject-verb and modifier-modified, signaled by affixation. Thus, for example, there is no formal similarity between the Spanish verb affix *-mos*, which marks the first person plural, and *nosotros/nosotras*, the pronoun to which it refers, or between *-s*, which marks the second person informal singular, and *tú*, the second person informal singular pronoun, but the agreement relationship serves to connect these two elements notwithstanding” (Bonta, 2018). In other words, agreement is a strategy for meaning association qua Thirdness, because it relies purely on conventional, symbolic relationships—relationships which, in languages with rich inflectional systems of agreement like Sanskrit, may be mediated across considerable morphosyntactic distances, minimizing reliance upon word order.

If the negative constraint [-3] conditions Chinese morphosyntax as we have shown it to do for the lexicon, we would certainly expect a minimization of agreement ([3]) as a strategy for meaning association, and this is indeed the case. Mandarin Chinese is as close an approximation to a pure analytic language as any attested. There is absolutely no subject-verb (or, for that matter, object-verb) agreement in Chinese. Moreover, there is little overt inflectional evidence of a copula or verb inflection; Chinese verbs exhibit not only no person or number, but also no inflectional tense or mood. The verb *shì*, ‘be, is, am, are, was, were’, is used only with subject complements. There are no irregular or suppletive verb forms. Typologically, the morphosyntax of Chinese is in diametric contrast to the richly-inflected Indo-

European languages (including even the comparative simplicity of modern English; Chinese learners of English are greatly vexed by the need for the third-person singular present tense *-s* and the past-tense *-ed*, for example).

We note also the near-total absence of an overt plural marker in Chinese, a typological oddity. The suffix *-men* distinguishes plural pronouns from their singular counterparts (e.g., *wǒ-men*, ‘we’ vs. *wǒ*, ‘I’), and may also be optionally deployed after nouns denoting persons (especially family members and children), but otherwise, no morphological distinction is made in Chinese between singular and plural. In other words, “plurality” is a notion accorded no morphosemiotic prominence—but this is not surprising in light of our contention that [-3] is a negative conditioning factor. Plurality, after all, is one of the most conspicuous phenomena associated with Thirdness. Thus, we take the lack of an overt plural marker for the overwhelming majority of Chinese nouns to be another consequence of the negative constraint [-3].

As for the positive constraint [+12], we consider the manner in which the (Peircean semiotic) Subject is morphosyntactically associated with the Predicate in Chinese, wherever the Subject consists of multiple arguments. Whenever the semiotic Subject consists of only a morphosyntactic subject (i.e., the verb is intransitive, and the subject is its only argument), the Subject-Predicate relationship is expressed via contiguity/[2], as in *wǒ kànjiàn*, ‘I see’, where subject/Subject *wǒ* precedes the verb/Predicate. Otherwise couched in linguistic terminology, subject incorporation is not admissible in Chinese.

However, whenever the semiotic Subject consists of a morphosyntactic subject and direct object (i.e., the verb is transitive, and the subject and direct object are its only two arguments; Peirce termed Subjects with two arguments dyads), the direct object is in some cases incorporated into the verb in a verb-object compound, and in others either precedes the verb (sometimes, though not always, accompanied by the preposing particle *bǎ*) or follows the verb, as a discrete entry. In the first case (i.e., of direct object incorporation), the relationship verb-direct object (or, in other words, the relationship between the Predicate and the part of the Subject corresponding to the patient or grammatical direct object) is accomplished via blending/[1], and in the second case (i.e., of discrete direct objects), it is accomplished via contiguity/[2]. In either case, the grammatical subject (i.e., the part of the Subject corresponding to the agent or grammatical subject) is kept discrete, as with intransitive verbs described previously. Thus, in instances where the direct object is incorporated into the verb, the

Subject-Predicate relationship is manifest as a structural dichotomy, with the subject-verb part represented via contiguity/[2], and the subject-direct object part represented via blending/[1]. Examples of each of these cases follow:

Case 1: Discrete subject, direct object incorporated (S VO):

wǒ *kàn-shū*
 I (subject) read-book (direct object)
 ‘I read the book/books/a book’

Case 2: Discrete subject, discrete preposed direct object (S O V):

wǒ *bǎ* *chuānghu* *dǎkāi*
 I (subject) [preposing particle] window (direct object) open
 ‘I open the window’

Case 3: Discrete subject, discrete postposed direct object (S V O):

wǒ *kànjiàn tāmen*
 I (subject) see them (direct object)
 ‘I see them’

In cases where the semiotic Subject consists of either 1) a subject and a second argument that is not a direct object or 2) more than two grammatical arguments (as, e.g., a subject, a direct object argument, and a locative argument), one (but not more than one) of the additional arguments besides the subject or direct object may be represented via a postverb, a sort of preposition-like verb (or verb-like preposition) that is affixed to the main verb, creating a species of compound verb that in effect incorporates the place or manner of the action into the grammatical predicate. Such a postverb is *zài*, meaning ‘[be] on, in’, and frequently incorporated as a postfix into verbs like *fàng*, ‘put, place’ (*fàngzài*, ‘put on, place on’) and *tú*, ‘smear’ (*túzài*, ‘smear on’). In some cases, postverb compounds may arise in circumstances where there is no direct object, but an additional argument, such as a locative, is called for. For example, the aforementioned *-zài* may be affixed to *zhù-*, ‘live’, to form *zhùzài*, ‘live in’. Another such postverb that requires a locative argument and often occurs in intransitive contexts is *-dào*, ‘[go] to, towards’, as in *bāndào*, ‘move to’.

Such postverbs and the compounds created therewith are a central characteristic

of Chinese morphosyntax, and are introduced early in Chinese pedagogy. Some examples of the uses of such postverbs follow:

Case 4: Discrete subject, no direct object, partially incorporated locational argument (S VL)

- a. *tāmen* *zhù-zài* *běijīng*
 they (subject) live-in Beijing (locative argument)
 ‘they live in Beijing’
- b. *wǒ* *bān-dào* *měiguó*
 I (subject) move-to America (destinational argument)
 ‘I move to America’

Case 5: Discrete subject and direct object, partially incorporated locational argument (S O VL)

- a. *wǒ* *bǎ* *zhège bāo* *dài-dào jiàoshì*
 I (subject) [preposing particle] this bag (direct object) bring-to classroom
 (locative argument)
 ‘I bring this bag to the classroom’
- b. *tā* *bǎ* *shū* *fàng-zai bēibāo*
 he (subject) [preposing particle] book (direct object) put-in backpack
 (destinational argument)
 ‘he puts the book in the backpack’

Recall that, as previously noted, the Peircean semiotic Subject consists of the entire set of logical arguments associated with the predicate; thus, the “Subject” of ‘John gives a book to Mary’ is the set {John, Mary, book}. In similar fashion, the Subject of the Chinese verb *fàng* in the second example under Case 5 preceding is the set {*tā*, *shū*, *bēibāo*}, since the notion of “putting” always involves a “putter”, a “thing which is put”, and a “location where it is put”. All of these three arguments are intrinsic to the meaning of put/*fàng*. It will be appreciated from the preceding examples that, in cases where the semiotic Subject may be resolved into three or more grammatical arguments, one argument other than the subject/agent and direct object/patient may be represented as partially incorporated into the Predicate by the blending of a locational or directional postverb with the main verb as a single

compound. Moreover, for intransitive verbs with a locational or directional argument, a postverb compound construction is typically deployed along with a discrete subject. Along with instances of discrete subjects paired with incorporated direct objects noted under Case 1 preceding, Cases 4 and 5 all involve representation of the semiotic Subject as a morphosyntactically iconic dichotomy whereof either a) the grammatical subject is a discrete entry and the direct object is a blended (incorporated) entry (Case 1), b) the grammatical subject is a discrete entry and the locational or directional argument a partially-blended entry as a postverb (Case 4) or c) the grammatical subject, direct object, and, potentially, other arguments (like the indirect object) may all be represented as discrete entries vis-à-vis the verb, while one additional argument (locational or directional) is represented as a partially-blended entry with the verb (Case 5). For the set of discrete arguments, the mode of representation, again, is contiguity/[2], and for the blended direct object or partially-blended locational or directional argument, blending/[1].

For greater clarity, we may represent the composite structure of the Chinese Subject with respect to the predicate as follows:

$$\begin{array}{l} \{S_n\} \text{ ____ } P \\ [C_n] \end{array}$$

Here, $\{S_n\}$ is the set of n terms corresponding to the semiotic Subject (for example, the subject and direct object), $[C_n]$ is the Categorical constraint corresponding to the mode of relationship with the Predicate of each of the n terms of the semiotic Subject, ____ represents the Subject in relation to the Predicate (the choice of notation motivated by Peirce’s characterization of the potential Subject as “blanks to be filled” [“New Elements”, Peirce 1998, pp. 310-311]), and P is the predicate. For example, Case 1 foregoing may be represented as follows:

$$\begin{array}{l} \{w\check{o}, sh\bar{u}\} \text{ ____ } k\grave{a}n \\ [2], [1] \end{array}$$

This diagram shows that the set of terms representing the Subject ($\{w\check{o}, sh\bar{u}\}$) are represented as connected to the predicate ($k\grave{a}n$) via a Secondness/[2] and a Firstness/[1], respectively. In cases such as this, the Subject is represented

morphosyntactically as a simple dichotomy [1], [2].⁶

On the other hand, Case 4a is represented as follows:

$$\{tāmen, běijīng\} \text{ ___ } zhù$$

$$[2], \dots [12]$$

The motivation for this representation is that, whereas the grammatical subject *tāmen* stands in a relationship of contiguity with the predicate *zhù*, the second argument, the location *běijīng*, stands in a morphosyntactic relationship consisting of both a discrete element (the noun *běijīng*) and a blended element (the indicator of spatial relationship *-zài*). In other words, on careful examination we observe that the Subject is resolved, in this case, not into a simple [1], [2] dichotomy but into a [12], [2] dichotomy, the choice of [12] for the locational term being motivated by its obvious additional resolution into a stronger ([2]) and weaker ([1]) morphosyntactic element.

For Case 5a, we have a more complex instance of a Subject composed of three separate elements, to wit:

$$\{wǒ, zhège bāo, jiàoshì\} \text{ ___ } dài$$

$$[2], [2], \quad [12]$$

In this instance, we see that the grammatical subject and direct object are represented as discrete elements with respect to the predicate ([2]), while the destinational argument *jiàoshì* is represented as a composite of the blended directional marker *-dào* and the discrete element *jiàoshì* ([12]). Note that, in this, as with other examples, the locational markers and prepositional particles do not figure in any significant way into the semiotic structures under consideration, at this level of analysis, though it is likely the case that they will also require mapping at subtler levels of semiotic representation. For now, our primary concern is the relationship between Subject and Predicate, for which explicit particles, postfixes, and the like, serve merely to make explicit semiotic information already inherent in the Predicate. Note also that the rule stated previously that, in the case of more than two arguments (beyond a subject and direct object; Peirce termed Subjects with more than two arguments polyads), one—but not more than one—additional argument may be represented as a partially-blended element via a postverb construction—amounts to a rule that, if there are

multiple arguments represented via the same Categorial constraint, that constraint will be [2], not [1] or [12]. For Case 5a shown above, this results in a double dichotomy by bifurcation, or in other words, the Subject is first resolved into a strong element (the subject and direct object, both [2]) and a weak element (the destinational argument, [12]), and the strong element of the Subject (the subject + direct object) is further resolved into a finer dichotomy, also consisting of a strong (subject) and weak (direct object) component. Recall that the admissibility of such bifurcations in [12] is described by Peirce as follows: “[A] genus characterized by Reaction will by the determination of its essential character split into two species, one a species where the Secondness is strong, the other a species where the Secondness is weak, and the strong species will subdivide into two that will be similarly related, without any corresponding subdivision of the species.” (CP 5: 69)

Thusly is the dichotomy of [+12] clearly manifest in Chinese morphosyntax insofar as the Subject and its relationship with the Predicate are concerned. That it is an unequal dichotomy consistent with the character of [12] previously established is also plain: [2]/contiguity is clearly the default strategy, as evidenced by the fact that the grammatical subject/agent is never incorporated into the verb, and, as noted previously, it is the only strategy that may be deployed with multiple elements within the same Subject. The direct object/patient may be incorporated, but if there are other arguments present beyond subject/agent and direct object/patient, only certain of those, and not the agent and patient, are susceptible to incorporation, and never more than one of them in a given proposition.

In addition to the Subject and its relationship with the Predicate, the Chinese predicate per se may be split into contrastive or complementary elements via a range of strategies, all of which have the effect of representing the Predicate as an unequal dichotomy emblematic of [12]. One of these strategies, familiar to every introductory-level Chinese student, is the contrastively-structured yes-no question, in which the verb is repeated, and the second instance is negated with one of two negative prefixes, *bú-* or *měi-*, as in *shì-bú-shì*, ‘Right?/Isn’t that so?’ (literally, ‘be-no-be’), or *yǒu-měi-yǒu*, ‘Is/are there?/Do you, they/Does he, she have?’ (literally, ‘There is-no-there is’). For example:

nǐ xǐhuan shàng hǎi, shì-bú-shì?
 you like Shanghai be-not-be

‘You like Shanghai, right?’

tā yǒu-méi-yǒu

he have-not-have

‘Does he have [it/them/any, etc.]?’

tāmen yào-bú-yào qián

they want-not-want money

‘Do they want money?’

Per Peirce, the affirmative and the negative are embodiments of Firstness and Secondness, respectively (CP 1: 359), and so the affirmative and negative in such dichotomies correspond to the weak and the strong elements of [12].⁷ We may represent such split predicates schematically according to the following formula:

$$\begin{array}{c} [C_{ij}] \\ (P_{ij}) \end{array}$$

where (P_{ij}) signifies the resolution of the Predicate into two contrastive elements i and j , and the superscript $[C_{ij}]$ represents the Categorical correspondence ([1] or [2]) for each element. In combination with our previous representation of Subjects, we represent a complete Subject-Predicate for such split predicates as:

$$\begin{array}{c} [C_{ij}] \\ \{S_n\} \text{ --- } (P_{ij}) \\ [C_n] \end{array}$$

We represent the above example

tāmen yào-bú-yào qián

as follows:

[1] / [2]
 {*tāmen, qián*} ___ (*yào/búyào*)
 [2] , [2]

This pattern of predicate-splitting is also used for experiential yes/no interrogatives; one common strategy is to use *méiyǒu*, ‘not-have’. For example:

nǐ qù-guó zhōngguó méiyǒu
 you go-EXPERIENTIAL China not-have
 ‘Have you been to China?’

or

nǐ yǒu méiyǒu qù-guó zhōngguó
 you have not-have go-EXPERIENTIAL China

We represent this sentence diagrammatically as follows:

[1] / [2]
 {*nǐ, zhōngguó*} ___ (*yǒu/méiyǒu, qùguó*)
 [2], [2]

Split predicates in affirmative sentences may be realized by a range of strategies. So-called “resultative compounds”, whereby a secondary or unmarked stative verb is conjoined with a main verb to form a compound, are extremely common in China, and are often preferred to equivalent periphrasis. The two members of such pairings are frequently separated by *-de-* (affirmative) and *-bu-* (negative). Some examples of these structures, also given in Bonta 2020, are shown following, with the stronger member of the pair shown in boldface:

tīng-de-dǒng
 hear-POSSESSIVE-understand
 ‘hear and understand’ (often used instead of ‘understand’ when comprehension of a spoken utterance is referenced)

tīng-bu-míngbái

hear-NEGATIVE-understand

‘don’t [hear and] understand’

shuō-bu-tài-dìng

say-NEGATIVE-too-be definite

‘cannot say too definitely/for sure’

kàn-bu-jiàn

see-NEGATIVE-catch sight of

‘cannot see (i.e., I cannot catch sight of)’

chī-bu-xià

eat-NEGATIVE-go down

‘cannot get down any more (food)’

The designation of the second element as the strong element is motivated by the fact that it is this element that can be negated. Such a construction would thus be rendered in our proposed notation as:

[1]/ [2]

(*kàn/bujiàn*)

Regarding such compounds, we have observed:

It should be plain that Chinese resultative compounds conform precisely to the pattern embodied by [12] as set forth by Peirce and clarified by Robertson: they involve the conceptual resolution of a single concept into two unequal parts, as with, e.g., *tīng-de-dǒng*, in which “understanding” is resolved into “hearing” (weak) and “understanding” (strong) or *kàn-bu-jiàn*, where “seeing” is resolved into “seeing” (weak) and “catching sight of” (strong). Unlike dilexic, disyllabic words, these compounds do not constitute lexemic entries per se; rather, the formation of resultative compounds is a productive morphosyntactic process with many thousands of potential configurations. (Bonta, 2020)

Further details concerning the formation of resultative compounds are also given in Bonta (2020).

Chinese also makes heavy use of adjectival predicates, as also described in great detail in Bonta (2020). For our purposes here, we simply note that adjectival predicates in Chinese can never appear without either a modifying adverb or aspectual particle. As a result, *hǎo*, ‘[be] good/well’, can never be used with a simple subject, as, e.g., **wǒ hǎo*, ‘I am good’. Instead, some adverb of degree, like *hěn*, ‘very’, or *fēicháng*, ‘really’, must be included, even if no strong degree is intended. Thus, *wǒ hěn hǎo* means either ‘I am good’ or ‘I am very good’. On the other hand, we can say *wǒ bu-hǎo*, ‘I am not good,’ or *wǒ hǎo-le*, ‘I am fine’ (where *-le* is a particle denoting completion or presentness), because in these two cases, the adjectival predicate is clearly bipartite, consisting in the first instance of negative adverbial particle *bu-* + adjectival predicate *hǎo*, and in the second, of adjectival predicate *hǎo* plus aspectual particle *-le*. In every instance, we see that Chinese requires adjectival predicates to consist of two elements, the adjectival predicate per se and an adverb or aspectual particle.

Here, too, we see the operation of the positive constraint [+12] configuring Chinese predicates. In such cases, it is the modifying particle (*hěn*, *bu*, or *le* in the aforementioned examples) that is serving as the weak (i.e., qualifying) element alongside the strong adjectival predicate:

$$\begin{array}{c} [1]/[2] \\ \{w\ddot{o}\} \text{ ___ } (h\ddot{e}n/h\ddot{a}o) \\ [2] \end{array}$$

To summarize, we have identified four morphosyntactic manifestations of [+12] in Chinese, one of which (the use of direct object incorporation and postverbs to split the Subject into two domains with different modes of association with the Predicate) is concerned with the Subject-Predicate relationship, and three of which (the use of affirmative-negative pairs for yes/no questions, the use of resultative compounds, and the requirement of a modifying element, whether adverb or aspectual particle, with adjectival predicates) are concerned with the structure of the predicate itself. Each of these four morphosyntactic manifestations of [+12] involve the construction of an unequal dichotomy. For additional clarity, they are shown together on Table 4:

Table 4. The unequal dichotomy [12] in Chinese morphosyntax

I. Subject Predicate Relational Dichotomy:

{wǒ, zhège bāo, jiàoshi} ___ dài
 [2], [2], [12]

II. Predicate-internal Structural Dichotomy Types:

[1] / [2]
 {tāmen, qián} ___ (yào/búyào)
 [2] , [2]

[1]/ [2]
 (kàn/bujiàn)

[1]/[2]
 {wǒ} ___ (hěn/hǎo)
 [2]

2.1.3 Chinese phonology

The realm of phonology, with its distinctive features and phonemes, has been relegated by many linguists and semioticians to asemiotic status; Hjelmslev’s meaningless “figurae”, of which phonemes and distinctive features were held to be prime examples, became the basis for the model of so-called double articulation of human language. But in the much more perceptive cadences of the Prague School tradition, the primordial components of language—distinctive features in particular—are indeed meaningful, albeit in an extremely inchoate sense, their semiotic force being derived mostly from “mere otherness” inherent in featural oppositions. Wrote Jakobson and Waugh:

The sole *signatum* of any distinctive feature in its primary, purely sense-discriminative role, is “otherness”: as a rule a change in one feature confronts us either with a word of another meaning or with a nonsensical group of sounds.... Distinctive oppositions have no positive content on the level of the *signatum* and announce only the nearly certain unlikeness of morphemes and words which differ in the distinctive features used. The opposition here lies not in the *signatum* but in the *signans*: phonic elements appear to be

polarized in order to be used for semantic purposes. Such a polarization is inseparably bound to the semiotic role of distinctive features. (Jakobson & Waugh, 1979, p. 47)

Phonemes, which are derived from the distinctive features, may be held to be more meaningful, i.e., less semiotically vague, while still falling far short of the assertive precision of propositions and other full-fledged instantiations of language. Yet if these primordial features of language are indeed semiotic, then we would expect the Categories to be manifest in their workings, just as has proven to be the case elsewhere. Having already shown evidence for the consistency of semiotic structuring in the lexicon, morphology, and syntax in Chinese, we predict that the conditioning constraints [-3] and [+12] are operative at the phonological level also. But in order to test this, we first need to consider in what ways the Categories might be manifest in phonology.

In general, phonology would appear to be a domain where Firstness preponderates. The distinctive features are in effect qualities of sound that shape phonemes and their allophones. But even here, phonemes per se are abstractions conventionalized by habits of mind, and so also embody Thirdness to a considerable degree; they are, in effect, very inchoate Symbols that give rise to Indexes (allophones), the phonetic existents that our physical senses perceive and interpret.

While languages display an enormous variety of phonemic possibilities, one universal contrast above all others divides phonemes into two broad domains: consonants and vowels. All languages have this contrast, and while every language has a different inventory, the essential contrasts between these two broad classes of sound are the same everywhere: whereas vowels tend to be continuous, pure voice, and very energetic, consonants (or at least, those that contrast maximally with vowels) tend to be voiceless, non-continuous, and reliant for their force on articulation rather than pure sound energy. In this way, the opposition between consonant and vowel generates syllables that the mind assembles into meaningful constituents.

Another contrast between vowels and consonants lies in their capacity to communicate meaning. In many Indo-European and Semitic languages particularly, consonants seem to bear most of the semiotic “load”, whereas vowels exist primarily to “color” words. That this is so is evidenced, among other things, by the fact that the abstractions we call “roots” in these languages are premised on consonants, not vowels, by the fact that so-called “phonaesthemes” (see discussion of these fascinating

entities in Jakobson & Waugh, 1979) depend for their force on consonantal, and not vocalic, configurations (such as the well-known fact that, in English, many words beginning with the consonant cluster *str-* denote long, thin objects), and by the fact that, in such languages, a passage with all its vowels deleted remains largely comprehensible, but not the reverse.

By contrast, vowels are frequently varied to denote changes in modality and tense, as with ablaut, or lengthened in speech to denote emphasis, incredulity, or some other affective coloring (as in, ‘*You did whaaaaat?*’).

Facts such as these suggest that, potentially at least, it is consonants that bear the primary capacity for symbolic representation/[3], whereas vowels are primarily affective/[1]. Consonants, after all, would seem to require greater mental effort to produce, involving, as they do, conscious manipulation of tongue, lips, teeth, etc., in conjunction with various areas of placement in the mouth, from the alveolar ridge to the velum, uvula, and even glottis. This is the reason that purely involuntary vocalizations (screams, gasps, groans, and the like) representing emotional states like pain, fear, pleasure, surprise, etc., tend to be vocalic—simple expulsions of breath across vibrating vocal cords. These are universal—a scream of pain or expression of surprise sound much the same across languages—and all are denotative of states that approximate Firstness.

However, this contrast between consonants and vowels, as to their semiotic freighting, is not at all evident in Chinese—and indeed, in many other language groups outside of the domain of Indo-European and Afro-Asiatic. In English, as we have already seen, the allowable number of syllables is much higher than in Chinese, and this is owing to the much greater number of allowable consonant configurations, including consonant clusters. Chinese, by contrast, allows few true consonant clusters, and those that it does allow typically involve a nasal or liquid + consonant, where nasals (/n/ and /m/) and liquid /r/ are among the least consonant-like of all consonants, and in many languages, are grouped with vowels. Moreover, consonant clusters within lexemes are not found, such that the canonical Chinese syllable, excluding the tone, is of the form (C)V(C), with only /n/, /r/, and /ŋ/ eligible as lexeme-final consonants. As a result, even across lexical boundaries, allowable sequences of C+C are extremely limited in Chinese. More complex consonant clusters of three entries, like *-spr-* and *-ndr-*, are completely excluded in all phonotactic environments.

This being the case, in contrast to English and many Indo-European languages, the

meaning of a passage of Chinese (written in pinyin or Romanized phonologic script) with vowels removed would, in all likelihood, be completely opaque.⁸ The absence of consonant complexes carrying the semiotic freight, therefore, may be characterized as the negative constraint [-3] operating at the phonological level (in this case, in the domain of phonotactic rules).

To discern the operation of the positive [+12] constraint at the level of Chinese phonology, we consider first how languages create meaningful sound, in general. The underlying operative principle both in the construction of phonemes from distinctive features and in the assembly of morphemes from phonemes is contrastive opposition. With phonemes, as we have already seen, the “mere otherness” of distinctive features provides for the differentiation of one phoneme from another; from the perspective of acoustic phonology, set forth by Jakobson and the Prague School, these features are represented in terms of paired opposites, such as compact/diffuse, grave/acute, and strident/mellow. The advantage of the acoustic approach is that such features may be realized in a variety of articulatory contexts; they represent a generalization and simplification of the very large ensemble of attested articulatory features.

Not only that, these opposing pairs give rise to phonemes which, in turn, rely upon another fundamental contrast—the opposition of vowel and consonant—to produce actual language. For European languages, these two layers of opposition suffice to create lexical and grammatical meaning, with other redundant features, including tone, constitutive of discourse modalities.

In Chinese, an additional dimension is required to complete the transition from inchoate sound structures to meaningful speech: sense-determinative tones. Every syllable in Chinese, as we have seen previously, is a lexeme, and therefore, every Chinese lexeme has one and only one vowel or diphthong, and a single corresponding tone.⁹ Unlike phonemes and their distinctive features, Chinese tones cannot be represented in terms of a matrix of oppositions. Their primordial meaning, as well as the sense in which they are applied to consonant-vowel configurations to instantiate lexemes, is not otherness but suchness. And this pairing of otherness with suchness is akin to pairing resistance/[2] with quality/[1] to achieve phonic completeness. As we have elsewhere suggested, this pairing is an unequal one, with the tone conceptually subordinate to the phoneme, as evidenced by the fact that tones may be lost in the second entry of disyllabic words, but not phonemes; by the fact that every Chinese lexeme perforce consists of an array of phonemes and distinctive features, but only

a single tone; and by the fact that in Chinese, as in all other attested languages, puns, rebuses, and the like rely for their significant force on phonemic similarity, and not on tonal similarity; and on the fact that, even in environments where tones cannot be discerned (songs, e.g.), words are still meaningful. Thus the structure of Chinese at the phonological level involves a dichotomy pairing the stronger, opposition-based operations of phonemes and their distinctive features with the weaker, qualitative modalities of the tones, [+12] embodied at the most elemental levels of linguistic signification.

What we have seen, then, is that the operation of the positive constraint [+12] and its negative concomitant [-3] are both present and pervasive at every level of Chinese that we have examined. Nor are these instances cherry-picked; they represent the entire range of Chinese typological distinctives, from tones to yes-no questions. While more thorough scrutiny of Chinese will of course yield exceptions, we anticipate that they will be far fewer than would ordinarily be the case with a language more susceptible to admixture with foreign influences (including semiotic imports). But with semiotics, we must not equate theoretical rigor with exceptionless results; we instead seek general, unmistakable trends consistently manifest in core, not peripheral, domains of a given semiotic landscape, and we are confident that, in the case of Chinese, we have found an abundance of such trends, all of which conform to a singular pair of primordial Categorical semiotic constraints.

To summarize, we have found the following evidence of the operation of the positive conditioning constraint [+12] and the negative conditioning constraint [-3] in Mandarin Chinese:

[+12]:

Morphosyntax:

- Multiple arguments (of complex or polyadic Subjects) related to Predicate via both contiguity and blending
- Predicates split according to pattern X-not X
- Verbs represented as double lexemes with *-de-/-bu-* interposed
- Adjectival predicates must include a second or complementary syllable, lexeme, or word, such as *bu-*, *-le*, or *hěn*.

Lexicon:

- Disyllabic/dilexemic words are the rule

Phonology:

- Pairing of “suchness”/[1] of tonal features with “otherness”/[2] of distinctive features

[-3]:

Morphosyntax:

- Absence of agreement morphology
- Near-absence of overt plural marker

Lexicon:

- All lexemes monosyllabic, with extremely high number of homophones owing to very limited set of allowable syllables

Phonology:

- Very limited inventory of consonants and consonant clusters

We now seek to adapt this method, *mutatis mutandis*, to several other languages in which semiotic structural constraints other than [+12] are prominent.

2.2 Sora and [1]

Pure Firstness/[1] as a semiotic structural constraint is often a characteristic of aboriginal cultures, as previously described in considerable detail elsewhere (Bonta, 2018). In general, we observed that so-called polysynthesis, the capacity of many aboriginal languages around the world to create “sentence-words”, is a morphosyntactic manifestation of [1]. That this is so can be envisaged in several different ways. For one thing, the merging of subject and predicate via the incorporation of grammatical subjects, direct objects, indirect objects, and other arguments of the predicate into a single indissoluble lexemic compound is typical of what we have elsewhere styled blending, namely, the strategy of idea association whereby subject and predicate (in the Peircean sense of the terms) are simply melded together into a whole. Such a strategy dispenses altogether with the Secondness of deixis and indexicality, which must be invoked wherever discrete lexemes are brought into mere juxtaposition. It is equally exclusionary of agreement, whereby meaningful association is achieved by inflectional morphology harmonizing one term with another (as with subject-verb agreement). Otherwise put, this strategy tends toward the representation of propositions as unitary morphosyntactic strings, in which the

complementarity of Subject and Predicate is not represented by any kind of structural separation.

We seek to discern whether, as with [12] in Chinese, [1] may be manifest as a parametrizing element in different linguistic layers within a single language. In light of the foregoing, a polysynthetic language is an obvious test subject.

Modern Asia is, compared with the Americas, New Guinea, and Australia, relatively impoverished insofar as polysynthetic languages are concerned (although strongly agglutinative languages, like Georgian and Tamil, are widespread). Eastern Siberia is an exception, although Siberian languages like Chukchi are nowadays spoken only by dwindling minorities, not unlike the vast majority of Native American and Australian Aboriginal languages.

Another pocket of polysynthetic languages is found in north central and northeastern India with the Munda languages, an offshoot of the Austro-Asiatic family. Not all Munda languages are polysynthetic; the most widely-spoken, Santali, which has between 7 and 8 million speakers, is strongly agglutinative. However, several other Munda languages are truly polysynthetic, and one of them, Sora, has the added advantages of being well-documented and (in relative terms) widely-spoken.

2.2.1 Sora morphosyntax

Sora, a language with over 400,000 speakers, is spoken mostly in the modern Indian states of Odisha and Andhra Pradesh. Although both Oriya (an Indo-Aryan language) and Telugu (a Dravidian language) are also widely spoken in areas inhabited by the Sora, the language has maintained much of its distinctive character, in no small part, perhaps, owing to the resilience of the Sora culture, including its shamanistic religion. In 1931, G. V. Ramamurti published a very complete grammar, lexicon, and textbook of the Sora language, in order to assist local officials in becoming familiar with what he characterized as an unusually difficult language.

First of all, Ramamurti's work leaves no doubt that Sora is a truly polysynthetic language; not only subjects and direct objects, but also indirect objects, instrumentals, adjectives, adverbs, and many other types of arguments are fully susceptible of incorporation into elaborate "sentence words", of which several examples are shown following:

nam- 'kid-t-am

catch-tiger-will-you

‘A tiger will catch you’

paŋ-ti- ‘da.r-ij- ‘te:n

bring-give-cooked rice-me-did

‘He brought and gave me cooked rice’

po:-puŋ- ‘kun-t-am

stab-belly-knife-will-you

‘I will stab you in the belly with a knife’

Such polysynthesis, or blending, is unquestionably the central and most conspicuous ordering principle of Sora morphosyntax. It means that, for Sora speakers, the default condition of any utterance is the blending of all elements into a single undifferentiated whole, precisely as Du Ponceau famously characterized polysynthesis two centuries ago.¹⁰ It does *not* mean, either for Sora or any other attested polysynthetic language, that concepts can never be expressed as discrete words; instead, discrete words tend to be deployed especially when emphasis—i.e., deixis—is needed, conforming precisely to the conditions of [2]. Per Peirce, nouns (or, at least, discrete nouns; it is unclear whether Peirce was aware of the phenomenon of polysynthesis) are strongly indexical (“Of Reasoning in General”, Peirce, 1998, pp. 17-18), and so their submersion within blended polysynthetic structures as the morphosyntactic norm may be taken as a predisposition to limit wherever possible the influence of Secondness, except where the exigencies of emphasis require some sort of deixis. Thus polysynthesis may be interpreted both as [+1] and [-2].

Moreover, in a true polysynthetic utterance (i.e., in which the finite verb, representing the Predicate, is blended with all of its arguments, representing the Subject, as in the preceding examples), the copula is likewise submerged. The grammatical copula embodies Peirce’s “sign of assertion”, the core or reifying Symbol of any proposition; as such it is not only a strong Thirdness/[3], but also, by serving to bring Subject and Predicate into relation with each other, involves an Index/[2]. From this perspective, it is seen how polysynthesis is characterized by [-3] and (doubly) by [-2]. Otherwise put, a polysynthetic assertion is as bereft as possible of overt representation of [2] and [3], inasmuch as it avoids any morphological instantiations

of discreteness ([-2]) or assertion ([-2] and [-3]).

We may provisionally represent a polysynthetic utterance as follows:

[1]
 {[*n*en], puŋ, ‘kun, am} ___ [‘po:]
 [1], [1], [1], [1]

‘po:-puŋ-‘kun-t-am

stab-belly-knife-will-you

‘I will stab you in the belly with a knife’ ([*n*en] represents non-explicit Sora first person singular *n*en, ‘I’)

2.2.2 Sora lexicon

The polysynthetic parameter in Sora triggers very significant and characteristic ordering at the lexical level, because Sora exhibits both “full” and “reduced” forms of all words susceptible to incorporation; or, as we more accurately style them, default/unmarked forms, which are used in incorporation, and enlarged/expanded/marked forms, which are deployed as discrete lexemes, for deixis or in otherwise marked circumstances. In Sora, many (though not all) nouns have dual forms, as do pronouns, adjectives and adverbs. Some common instances of such lexeme pairings are shown following, in Table 5:

Table 5. Long (independent) and short (incorporated) forms of Sora nouns

<u>Long</u>	<u>Short</u>	<u>Meaning</u>
<i>‘kina</i>	<i>‘kid</i>	tiger
<i>‘ali:</i>	<i>sa:l</i>	liquor
<i>amən</i>	<i>am</i>	you
<i>u:da:n</i>	<i>u:lən</i>	mango tree
<i>‘kudin</i>	<i>‘kunən</i>	a long knife
<i>‘kumulən</i>	<i>kumən</i>	rat
<i>jabmulən</i>	<i>jabən</i>	seed

Short forms such as those given above are never found outside of bound contexts, as incorporated arguments of Sora predicates. They are thus not true “words” in the sense that the long forms are; they are instead bound morphemes representing their

respective Objects as grammaticalized fragments of longer “sentence-words”. As we have already noted, nouns are highly indexical (especially as independent words). By transforming noun-words into noun-bound morphemes, as it were, Sora (and other polysynthetic languages) strive in effect to strip nouns of their indexical properties, or, in semiotic terms, to represent iconically that such a stripping is taking place at the level of semiotic structure. In other words, the creation of “bound nominals” is a manifestation of the negative constraint [-2] at the lexical level.

As for the positive constraint [+1], it will be readily appreciated that polysynthesis completely transforms the very notion of lexicon, since every “sentence word” produced via polysynthesis is, in some sense, a lexical entry, albeit, in many instances, a novel one. The reader may object that the notion of “word” is too subjective, and that there is no way to know how or whether speakers of polysynthetic languages (especially non-written languages, which constitute the vast majority of such languages) identify such things as “words” or “word boundaries”. To this objection Mithun made the following observations:

One might wonder why such long strings are considered single words at all, rather than phrases or sentences. Numerous criteria exist for determining word boundaries. The most important is that speakers generally know where a words begins and ends, whether they write their language or not. Second, speakers may pause between words, but not in the middle of them. An English speaker might say, “Speakers ... pause between words”, or “Speakers pause ... between words”, but not “Speak...ers pause between words”, nor “Speakers pause be ... tween words”. In many languages, each word, no matter how long, has only one primary stress. In some, the location of the stress is even quite predictable. ... In these languages, of course, it is especially easy to determine where words begin and end. There are other criteria for determining word boundaries as well. All together, they provide ample evidence that words in polysynthetic languages really are single words. (Mithun, 1983, p. 221)

Thus the reality of polysynthetic strings as words is non-controversial.

Mithun also notes that, among traditional Native American communities, the ability to neologize via polysynthesis—learned very late by L1 learners—is highly prized and admired:

Speakers of many American Indian languages often make a comment that speakers of English rarely seem to make. They are aware of who speaks their language especially well.... The speakers who are thus recognized generally share a specific trait: they frequently create new, morphologically complex words. They incorporate very often. They expand the lexicon by putting together new combinations. They actualize the possibilities their language has to offer. (Mithun, 1983, p. 235)

In other words, for speakers of polysynthetic languages, the lexicon is a fresh, ever-changing, living, breathing thing, whereof neology, novelty, and almost infinite variety are the cardinal attributes. Freshness, newness, and variety are, not coincidentally, attributes of Firstness. The lexicon of a polysynthetic language (or at least, that dominant portion of the lexicon encompassing polysynthetic strings in their nearly infinite potentialities) is the very essence of Firstness, manifest both syntagmatically and paradigmatically in the lexical realm.

2.2.3 Sora phonology

At the level of phonology, the operation of Categorical constraints may be harder to discern, because of the semiotic vagueness of phonemes and other sound elements. Of the two broad classes of phoneme common to all spoken languages, vowels and consonants, recall that their characteristics and modes of operation are essentially complementary, as discussed in the previous section. Whereas the most archetypal consonants—obstruents—are bounded, vowels are characteristically continuous or unbounded. Consonants may accomplish their task without voicing, whereas voiceless vowels, though not unknown, are exceedingly rare.

In acoustical terms, the bounded nature of consonants gives them their significative force; there are of course consonants that are continuous, voiced, etc.—nasals, sibilants, and liquids come immediately to mind—but these are ipso facto more vowel-like. The difference between two “pure” consonants, such as, e.g., /p/ and /t/, is not to be found in any positive quality like voice, but by the purely oppositional character of the distinctive features that define and demarcate them one from another. As discussed previously, consonants in many languages, including English, are susceptible to clustering and to forming sequences that are readily associated with meaning, even in the absence of vowels, as every Semitic or Indo-European linguist is well aware; with a few exceptions, consonants provide the scaffolding for root

formation in such languages, and hence for most of significant force of words.

Consonants are therefore both highly indexical/[2] and symbolic/[3], especially in aggregates—much more so than vowels. The latter, for their part, are much more iconic/[1], providing qualitative coloring for language in the same way that spectral colors do for the visible realm, or feelings for the conscious. To remind the reader of what we have previously observed concerning vowels and their properties, the iconicity/Firstness of vowels may plainly be seen by considering that vowels are the most characteristic elements of onomatopoeic words, which mimic sounds; that vowels, or very vowel-like consonants, are used to represent purely emotional responses, such as [a] (“Ah!”), [o] (“Oh!”), [u] (“Ooh!”), “Shhh!”, etc.; and that vowel variation is frequently used to denote different “shades of meaning”, as with verb tense, voice, and mood, noun number (‘mouse’ vs. ‘mice’, e.g.), and even the passage from verb to noun to adjective, as is frequently accomplished in Sanskrit via the so-called *guna* and *vrddhi* grades of root vowel augmentation.

All of which is to say that the relationship between consonants and vowels is akin to that between both Seconds and Thirds, on the one hand, and Firsts on the other (although of course, as is always the case in the world of experience, the lines between the two are often blurred or overlapping, as with vowel-like consonants and consonant-like vowels. But this should not distract from the essential semiotic truth of the matter).

We therefore would expect that, for a language to manifest the lens of [1] at the phonemic level, vowels will be distinctively prominent or emphasized with respect to consonants. This may possibly be accomplished in various ways (having words that are much more vowel-heavy than consonant-heavy, for example), but in the case of Sora, this trait is immediately and unmistakably manifest upon inspection of the phonemic inventory: Sora exhibits an extraordinary number of phonemic vowels and diphthongs. Every vowel except [ə] may be short, half-long, or long, all three varieties of which are fully phonemic (Ramamurti, 1931, p. 6), yielding a total of 25 different vowels. Added to that the “remarkably large number of diphthongs and triphthongs” noted by Ramamurti (20 in all, but the real number is apparently much larger, because he conflates diphthongs involving vowels of different length, which are presumably just as distinct as the vowels themselves; for example, under diphthong [ai], he gives examples of both [ai] and [a:i]. As a result, the true number of distinct diphthongs may be in excess of 40, but the data given is not sufficient to know for sure), Sora

presents as rich and varied an inventory of vowels as will be found among any of the world's languages—this in contrast to a decidedly limited inventory of consonants (16 are given by Ramamurti, including the glottal check). Tables 6, 7, and 8 display the vowels, diphthongs, and consonants of Sora:

Table 6. Sora vowels (superscripted “.” denotes mid-length)

i	i.	i:
u	u.	u:
ɪ	ɪ.	ɪ:
ʊ	ʊ.	ʊ:
e	e.	e:
ə		
o	o.	o:
ü	ü.	ü:
ö	ö.	ö:

Table 7. Sora diphthongs and triphthongs (length variation not shown for diphthongs)

ai	au		
ia	ie	io	iu
ea	ei		
oa	oi	ou	
ua	ui	ue	
ua	ui		
ia:i			
e:ai	e:io		
ua:i			

Table 8. Sora consonants

k	g	ŋ
ʃ	j	ɲ
t	d	n
p	b	m
l	r	ʎ
s		
ʔ		

From these tables, the strong phonemic precedence of vowels over consonants in Sora is very evident. We propose that this notable asymmetry is a consequence of the lens of Firstness/[1] at the phonemic level. In particular, the extraordinary luxuriance of vowels and diphthongs may be represented as the positive constraint [+1], and the comparative paucity of consonants as the negative constraint [-2].

In effect, the most distinctive characteristics of Sora morphosyntax, lexicon, and phonology are seen, in an admittedly preliminary way, to display a very clear common structural constraint, which we have identified as [+1]. Were our acquaintance with this little-studied language more thorough, we anticipate that more such instances of the operation of this constraint would be evident, as we have found them to be in Chinese.

The study of polysynthetic languages presents many vexing concerns for the investigator, not the least of which is that such languages are nearly impossible for outsiders to learn. The ability to create polysynthetic neologisms is acquired very late by L1 learners, according to Mithun (1983, pp. 235-236), and this skill is easily lost once such languages come in contact with non-polysynthetic colonizing idioms like English and French. Moreover, for many such languages, polysynthesis is merely a stylistic option for certain circumstances, representing only a semiotic module, as it were, within a much more complex semiotic panorama. Thus, while a very few polysynthetic languages (like Inuktitut and West Greenlandic) are essentially “pure”, the majority—including Sora—are more accurately styled as polysynthetic and agglutinating. Sora, as we have seen, has contracted forms for all pronouns and many nouns, but those nouns for which no contracted form is available cannot be incorporated. For these, as for all other non-contracted noun forms, an elaborate system of agglutinative case endings is available. Thus polysynthesis and related phenomena shown previously may be subsumed under the constraint [+1], but this constraint is typically found alongside others. However aesthetically appealing the notion of a language grounded entirely in Berkeleyan idealism (along the lines of, e.g., Borges’ “conjunctural Ursprache” of Tlön), the tyranny of objective reality forces recourse to some measure of indexicality. The constraint [+1] represents an ideal towards which certain languages, like Sora, clearly incline, but is insufficient to all of the semiotic demands of human language.

In contrast with [+12], the constraint [+1] seeks to blend all into one, to blur or erase all boundaries, to submerge the deixis of Secondness within an undifferentiated

whole. With polysynthesis, this means not only the blurring of morphosyntactic discreteness, it also means the dissolution of the distinction between sentence and word, between morphosyntax and lexicon.

We summarize the semiotic structures of Sora conditioned by the positive constraint [+1] as follows:

[+1]

Morphosyntax:

- Formation of polysynthetic “word sentences”, with incorporation of all arguments

Lexicon:

- Spontaneous expansion of the lexicon via polysynthetic neologizing
- Existence of both incorporated and free root forms corresponding to many nouns and pronouns

Phonology:

- Enormous variety of phonemic vowels and of allowable diphthongs and triphthongs

2.3 Tamil and [13]

Tamil, the oldest-attested and most conservative of the Dravidian languages, is a South Asian language that—unlike Telugu, Malayalam, and Kannada, the other three major literary Dravidian languages—has resisted the importation of Indo-Aryan to a considerable degree. Tamil alone has not adopted the characteristic Indo-Aryan aspirated stops, and while it has imported large numbers of Sanskrit words, it has preserved their Dravidian equivalents, reserving the former primarily for literary use. For our purposes, Tamil has the added advantages of a long and well-attested literary tradition and extensive documentation of its many dialects.

We propose that Tamil (and, in more diluted form, the other Dravidian languages) is conditioned by the positive semiotic constraint [+13] (Firstness of Thirdness), as well as a concomitant negative constraint [-2]. Like Firstness of Secondness/[12], Firstness of Thirdness (also sometimes called “Thirdness degenerate in the second degree” in Peirceana) is a “degenerate” Category, which Peirce characterized as follows:

The most degenerate Thirdness is where we conceive a mere Quality of Feeling, or Firstness, to represent itself to itself as Representation. Such, for example, would be Pure Self-Consciousness, which might be roughly described as a mere feeling that has a dark instinct of being a germ of thought. This sounds nonsensical, I grant. Yet something can be done toward rendering it comprehensible.

I remember a lady's averring that her father had heard a minister, of what complexion she did not say, open a prayer as follows: "O Thou, All-Sufficient, Self-Sufficient, Insufficient God." Now pure Self-consciousness is Self-sufficient, and if it is also regarded as All-sufficient, it would seem to follow that it must be Insufficient. I ought to apologize for introducing such Buffoonery into serious lectures. I do so because I seriously believe that a bit of fun helps thought and tends to keep it pragmatistical.

Imagine that upon the soil of a country, that has a single boundary line thus \circ , and not $\circ\circ$, or \odot , there lies a map of that same country. This map may distort the different provinces of the country to any extent. But I shall suppose that it represents every part of the country that has a single boundary, by a part of the map that has a single boundary, that every part is represented as bounded by such parts as it really is bounded by, that every point of the country is represented by a single point of the map, and that every point of the map represents a single point in the country. Let us further suppose that this map is infinitely minute in its representation so that there is no speck on any grain of sand in the country that could not be seen represented upon the map if we were to examine it under a sufficiently high magnifying power. Since, then, everything on the soil of the country is shown on the map, and since the map lies on the soil of the country, the map itself will be portrayed in the map, and in this map of the map everything on the soil of the country can be discerned, including the map itself with the map of the map within its boundary. Thus there will be within the map, a map of the map, and within that, a map of the map of the map, and so on ad *infinitum*. These maps being each within the preceding ones of the series, there will be a point contained in all of them, and this will be the map of itself. Each map which directly or indirectly represents the country is itself mapped in the next; *i.e.*, in the next [it] is represented to be a map of the country. In other words each map is *interpreted* as such in the next. We may therefore say that each is a representation of the country to the next map; and that point that is in all the maps is in itself the representation of nothing but itself and to nothing but itself. It is therefore the precise analogue of pure self-consciousness. As such it is *self-sufficient*. It is saved from being insufficient, that is as no representation at all, by the circumstance that it is not *all-sufficient*, that is, is not a

complete representation but is only a point upon a continuous map. (CP 5: 71)

Because of Peirce's comparative reticence on the properties of the degenerate Categories (which do not appear to have interested him as much as pure Firstness, Secondness, and Thirdness), we are left to elaborate on them as best we can. Previously, we observed the following, regarding the properties and manifestations of [13]:

Peirce's example of infinite self-representation via embedded maps is—like nestled Russian dolls and infinite reflections in facing mirrors—an instance of recursion which, in its turn, encompasses such notions as cyclicity and reproduction. What is a life cycle, after all, if not the process by which an organism represents itself in its offspring? And while no offspring is identical to its parent, it is nonetheless an iconic approximation thereto. The orbits and revolutions of a planet produce seasons and noctidial periods that, notwithstanding no particular summer is exactly like another, nor any day like any preceding or following, are icons of one another nonetheless. So it is with recursion, cyclicity, and reproduction generally: they (and hence, Firstness of Thirdness in general) will beget variation upon a theme, diversity arising from similarity—an “irreducible idea of Plurality”, indeed, but one always characterized by variety and freshness. (Bonta, 2015)

In other words, recursion involving variation on a theme is a conspicuous manifestation of [13]. To test for the constraint [+13], as with [+12] in Chinese, we must examine the most important typological traits of Tamil. As for the concomitant negative constraint [-2], we would expect to observe grammatical features tending to diminish opposition, complementarity, and “Otherness” in general.

2.3.1 Tamil morphosyntax

The most conspicuous of all Tamil grammatical traits is agglutination, the tendency to add multiple affixes to a stem for both verb conjugation and noun declension. Agglutination in Tamil does not extend to the incorporation of arguments, a crucial distinction between agglutination even in its most elaborate expressions, and true polysynthesis. Rather, agglutination is confined primarily to morphology appertaining to the part of speech represented by the stem; finite verbs do not incorporate nouns, adjectives, adverbs, etc., but typically only morphology representing properties intrinsic to the predicate. Similarly, noun affixes are limited to properties intrinsic

to nouns, like case and number. Among agglutinative languages, Tamil stands out as being among the most elaborate, allowing for up to eight or nine affixes in the formation of predicates.

At first blush, a verb paradigm in an agglutinative language like Tamil might appear similar to a typical conjugated verb in an Indo-European language like Spanish, Latin, or even Sanskrit. These latter have bedeviled generations of reluctant schoolchildren, in Europe and India alike, with their exquisitely-demarcated verb classes and stately, nuanced paradigms—not to mention their many vexingly irregular verbs. Tamil exhibits all of these tendencies, to be sure. Tamil verbs are conventionally divided into seven classes, and include some irregular forms (although nothing as outré as the more extreme suppletive irregular verbs found in Indo-European). Tense and many other verb properties are signaled via inflection.

But whereas the most elaborate Indo-European verbs might exhibit upwards of 70 or 80 distinct forms across all persons, numbers, tenses, aspects, moods, etc., a typical Tamil verb may exhibit hundreds of different forms. This is because, in an agglutinative language like Tamil, no verb property is ever expressed analytically. Deontic and epistemic modalities, as well as forms signifying dubiety, anger, emphasis, and other affective states, are all alike represented affixally.

For example, consider the Tamil root *cey-*, ‘do, make’. The simple present tense of this verb gives rise to at least eleven different forms, corresponding to a first person singular, a second person singular informal, a second person singular formal, a third person singular masculine, a third person singular feminine, a third person singular formal, a third person singular irrational, a first person plural, a second person plural, a third person plural rational, and a third person plural irrational. The fact that these forms are somewhat condensed, in informal speech, into nine or ten in no wise detracts from the luxuriance of this paradigm, relative to even the most elaborate Indo-European simple present tense (nine forms, for Sanskrit, with the dual number supplying an additional three to the conventional six forms found in Latin, Greek, and other languages). Rounding the Tamil number to ten, we find a similar number in both the past and future—noting that these forms are only for the affirmative. Negative verb conjugations, discussed elsewhere, will in the most conservative Tamil dialects double this number again, to roughly 60 forms. Every one of these forms can also be manifest by using various so-called “indicative auxiliaries” to form a wide array of compound tenses such as the present, past, and future perfect (all of which are affixal

in Tamil, involving the affixation of *iru-*, ‘be’, + the standard array of person and number morphology to the participial form of the main verb, rather than analytical, as they are in English, German, and the modern Romance languages). The basic 60 tense forms can also be represented in the present, past, and future progressive, also affixal (in which the form *-kkoṇṭu-iru-* + tense, person and number morphology are affixed to the participial stem). The verb can also be given a sense of reflexivity or “doing X for the grammatical subject” by affixing *-kkoṇṭu-* + tense, person and number morphology to the participial stem, or the sense of “doing X for someone else” by affixing *-kkoṭu-* + tense, person and number morphology to the participial stem. Each of these will yield approximately 60 additional forms apiece, for a rough total, so far, of about 300 forms. And there are a number of other such indicative auxiliaries, furnishing hundreds of additional forms. Each of these hundreds of forms can be made dubitative by the addition of a word-final affix *-ō*, emphatic by the addition of word-final *-ē*, reportative by the addition of word-final *-ām*, or interrogative by word-final *-ā*, for a total of well over a thousand potential forms. To these we may add hundreds more potential forms yielded by the passive, formed by the infinitive stem plus the experiential verb *-ppaṭu-* in its many instantiations of tense, person, and number. Some idea of this immense variety may be gleaned from Table 9.

Table 9. Tamil verb forms

cey-, ‘do, make’

Simple Present Tense (-*kir-*; 10 forms):

	Singular	Plural
1st	<i>cey-kir-ēn</i>	<i>cey-kir-om</i>
2nd	<i>cey-kir-āy</i>	<i>cey-kir-īrkaḷ</i>
3rd Masculine	<i>cey-kir-ān</i>	<i>cey-kir-ārkaḷ</i>
3rd Feminine	<i>cey-kir-āl</i>	“”
3rd Formal	<i>cey-kir-ār</i>	“”
3rd Irrational	<i>cey-kir-atu</i>	<i>cey-kinr-ana</i>

Simple Past Tense (-*t-*; 10 forms):

1st Sing. *cey-t-ēn*, etc.

Simple Future Tense (-v-; 9 forms):

1st Sing. *cey-v-ēn*, etc. (3rd irrational, singular and plural, is *ceyy-um*)

Literary Negative (stem plus personal endings; 10 forms, used for all tenses):

1st Sing. *ceyy-ēn*, etc.

Future Negative (infinitive stem + -māṭṭ-; 9 forms):

1st Sing. *ceyya-māṭṭ-ēn*, etc. (3rd irrational, singular and plural, is *ceyy-ātu*)

Present Progressive (past participle + -kkoṇṭ-iru-kkir-; 10 forms):

1st Sing. *ceytu-kkoṇṭ-iru-kkir-ēn*, etc.

Past Progressive (past participle + -kkoṇṭ-iru-nt-; 10 forms):

1st Sing. *ceytu-kkoṇṭ-iru-nt-ēn*, etc.

Future Progressive (past participle + -kkoṇṭ-iru-pp-; 9 forms):

1st Sing. *ceytu-kkoṇṭ-iru-pp-ēn*, etc.

Present Perfect (past participle stem + -iru-kkir-; 10 forms):

1st Sing. *ceyt-iru-kkir-ēn*, etc.

Past Perfect (past participle stem + -iru-nt-; 10 forms):

1st Sing. *ceyt-iru-nt-ēn*, etc.

Future Perfect (past participle stem + -iru-pp-; 9 forms):

1st Sing. *ceyt-iru-pp-ēn*, etc.

Disjunctive (past participle form + -viṭu-kir- [present tense, 10 forms]/ + -viṭ-ṭ- [past tense, 10 forms]/ + -viṭu-v- [future tense, 9 forms]):

1st Sing. Present *ceytu-viṭu-kir-ēn*, etc.

Progressive Disjunctive (past participle + -kkoṇṭ-iruntu-viṭu-kir- [present tense, 10 forms]/ + -kkoṇṭ-iruntu-viṭ-ṭ- [past tense, 10 forms]/ + -kkoṇṭ-iruntu-viṭu-v- [future tense, 9 forms]):

1st Sing. Present *ceytu-kkoṇṭ-iruntu-viṭu-kir-ēn*, etc.

To express action performed on behalf of someone else (past participle stem + -kkoṭu-kkir- [present tense, 10 forms]/ + -kkoṭu-tt- [past tense, 10 forms]/ + -kkoṭu-pp- [future

tense, 9 forms]:

1st Sing. Present *ceytu-kkoṭu-kkir-ēn*, etc.

Reflexive or action performed on behalf of or to oneself (past participle stem + -kkoḷ-kir- [present tense, 10 forms]/ + -kkoṇ-ṭ- [past tense, 10 forms]/ + -kkoḷ-v- [future tense, 9 forms]:

1st Sing. Present *ceytu-kkoḷ-kir-ēn*, etc.

Additional indicative auxiliary forms:

Continuous (“Keep X-ing”; past participle stem + *-kkoṇṭu* + *va* + tense + person/number [29 forms])

Alternate Progressive (past participle stem + *-kkoṇṭu-* + *-kiṭa-* + tense + person/number [29 forms])

Alternate Continuous (“Go on X-ing”; past participle stem + *-kkoṇṭu-* + *-ppo-* + tense + person/number [29 forms])

To express inability to do X (past participle stem + *-kkoṇṭu-* + *-kkiṭi-* + tense + person/number [29 forms])

To express “doing X slowly” (past participle stem + *-kkoṇṭu-* + *-vaḷi-* + tense + person/number [29 forms])

Emphatic forms (available for all verb forms using auxiliary *-kkoṇṭu-*; add emphatic *-ē-* to *-kkoṇṭu-* to form *-kkoṇṭē-* + remaining morphology as indicated above)

Interrogative forms (available for all verb forms given, formed by adding postclitic *-ā*)

Dubitative forms (available for all verb forms given, formed by adding postclitic *-ō*)

Passive (available for all preceding forms: transform final indicative auxiliary into infinitive stem and add *-ppaṭu-* + tense + person/number)

The total number of forms supplied by this table, either explicitly or by implication, is close to 2800, and there are many hundreds of others that may be generated from additional indicative auxiliary combinations available. It is important to emphasize, firstly, that each of these thousands of possible forms represents, not full sentences, as in the case of true polysynthesis, nor the tidy simple tense and mood-based paradigms of Indo-European. Nor—in contrast to the nearly infinite potentialities of true polysynthesis—are they neologisms; all are predictable entries in an extraordinarily luxuriant paradigm. Each of these forms is a verb, and represents one of a myriad possible variations on but a single theme—the verb itself (in the case of our example,

cey-, including its infinitival [*ceyya-*] and past participial [*ceytu-*] versions).

Less luxuriant but withal impressive is the Tamil noun case system, where the eight cases of Indo-European are exceeded by Tamil's fifteen or more allowable case affixes (including at least three different concurrent genitive forms). Although Tamil, unlike Sanskrit, has no dual number, the Sanskrit dual usually only displays three distinct forms across eight cases, as well as a maximum of seven plural forms, for a maximum total of 18 forms for Sanskrit nouns (although 15 or 16 distinct forms is more usual). Tamil, by contrast, has a plural affix *-kaḷ*, to which are added all of the case endings available for the singular (or, in some cases, as with the Genitive 3 plural *-inatu* listed below, a modified version of the singular affix [*-atu*] for use in the plural context), yielding a bare minimum of 30 possible declensional forms. But to these forms may also be affixed postclitics like concessive/dubitative *-ō*, conjunctive *-um*, interrogative *-ā*, emphatic *-ē*, and reportative *-ām*, among others, yielding manifold more possibilities. A sample paradigm is shown in Table 10.

Table 10. Tamil noun forms (manitan, 'man')

	Singular	Plural
Nominative	<i>manitan</i>	<i>manitar-kaḷ</i>
Genitive 1	<i>manitan-uṭaiya</i>	<i>manitar-kaḷ-uṭaiya</i>
Genitive 2	<i>manitan-in</i>	<i>manitar-kaḷ-in</i>
Genitive 3	<i>manitan-atu</i>	<i>manitar-kaḷ-inatu</i>
Dative	<i>manitan-ukku</i>	<i>manitar-kaḷ-ukku</i>
Accusative	<i>manitan-ai</i>	<i>manitar-kaḷ-ai</i>
Locative 1	<i>manitan-il</i>	<i>manitar-kaḷ-il</i>
Locative 2	<i>manitan-iṭam</i>	<i>manitar-kaḷ-iṭam</i>
Locative 3	<i>manitan-iṭattil</i>	<i>manitar-kaḷ-iṭattil</i>
Sociative	<i>manitan-ōṭu</i>	<i>manitar-kaḷ-ōṭu</i>
Vocative	<i>manitan-ē</i>	<i>manitar-kaḷ-ē</i>
Ablative 1	<i>manitan-iliruntu</i>	<i>manitar-kaḷ-iliruntu</i>
Ablative 2	<i>manitan-iṭattiliruntu</i>	<i>manitar-kaḷ-iṭattiliruntu</i>
Ablative 3	<i>manitan-inru</i>	<i>manitar-kaḷ-inru</i>
Instrumental	<i>manitan-āl</i>	<i>manitar-kaḷ-āl</i>

Such affixal luxuriance far exceeds the vexing but learnable patterns of Indo-European, while falling short of the nearly infinite neologistic variety allowable by polysynthesis. Note also that, whereas in an Indo-European paradigm, the emphasis is

ultimately on the theme, as emblemized by the notion of roots and thematic vowels, in Tamil, the emphasis is laid on the variation, as emblemized, for verbs, by the formally inconspicuous stem with up to eight or nine affixes concatenated thereto.

As we have seen, variation on a theme is a token of [13] *par excellence*. The paradigm, *sensu lato*, is the embodiment of this notion. Hence we may say that agglutination, giving rise as it does to maximal paradigmatic variation, is the primary morphosyntactic manifestation of this Category, lending considerable credence to our initial hypothesis, that [+13] is the primordial positive constraint conditioning Tamil.

A byproduct of agglutination is the near-total absence of periphrasis. Periphrasis is merely the realization of the grammatical predicate via the juxtaposition of discrete elements (such as with the English future tense using *will* + verb instead of affixation), the strategy of meaning relation corresponding to Secondness. The absence of periphrasis in Tamil is therefore a clear manifestation of the negative constraint [-2] in Tamil morphosyntax.

There are other manifestations of [-2] in Tamil besides. As mentioned previously, Tamil conjunctions occur as postclitics; in particular, there is no unbound lexeme signifying ‘and’. Instead, postclitic *-um*, ‘and’, is affixed to each noun or pronoun in a conjoined pair or group. As a result, the morphosyntactic representation of a conjoined pair, instead of having the form ‘A and B’, is represented as ‘A-*um* B-*um*’. For example, in Tamil we might see:

nān-um nī-um

I-and you-and

‘you and I’

paravaiy-um mīn-um

bird-and fish-and

‘the/a bird and the/a fish’

Affixed to a single noun or pronoun, *-um* also means ‘also, too’, as in *avan-um pō-kir-ān*, ‘he too is going’. In this way, the contrastive morphosyntactic character of the oppositional or complementary idea of ‘A and B’ is rendered as formally innocuous as possible. Instead of a stark morphosyntactic dyad whose complementarity is highlighted by a single unbound morpheme (‘and’ or its lexemic equivalent) standing

amidmost, the notion of conjunction is represented by bound postclitics, one for each entry. Iconically, the *A-um B-um* construction represents conjunction, in general, not as an explicit, standalone Index (a word like ‘*and*’), but as a subordinate, incidental Quality bound to each of a pair or series of entries—concatenated in some relational sense, to be sure, but not consisting as a pair or pairs whereof the “pairedness” is independently represented.

In a similar way, conjoined independent clauses, of the form ‘he did X and did Y’ are disallowed in Tamil (and, indeed, in most of the South Asian Sprachbund). Instead, all but the final verb are represented by past perfect participles (usually called “converbials” or “conjunctive participles”), with the final verb alone being inflected; the literal equivalent in English of such a construction is something like ‘having done X, he did Y’. This mode of conjoining predicates can be extended to an indefinite number of entries, as, e.g., ‘having done W, having done X, having done Y, he did Z’ (in other words, ‘he did W and he did X and he did Y and he did Z’). A couple of examples illustrating this construction follow:

avan vīṭṭ-ukku *pōy* *cāppi-ṭṭ-ān*
 he home-DATIVE having gone eat-PAST-3rd Sing. Masc.
 ‘He went home and ate (lit., ‘Having gone home, he ate’)’

avan vēlai ceytu, *vilaiyāṭṭi,* *tūṅku-kir-ān*
 he work having done having played sleep-PRES-3rd Sing.
 ‘He works, plays, and sleeps (lit., ‘Having worked, having played, he sleeps’)’

In such instances—which represent the default morphosyntactic option for conjoining independent clauses (or rather, the Tamil equivalents of conjoined independent clauses in English)—the overt conjunction is done away with altogether, and the associated Sign of Assertion, signaled by finite verb inflection, is reduced to a single instance (the final verb in the series), the other actions represented, not as free-standing predicates independent from, and complementing one another, but as states or qualities subordinate to a single, primary predicational assertion. Consider, by contrast, the English equivalent, whereof the (potentially repeated) overt conjunction “**and**” and finite verb forms all place iconic semiotic emphasis on the independent, free-standing, and ultimately Indexical character of each of the verbs: “he works [**and**]

plays **and** sleeps”.

Overall, the notion of conjunction involves strong Secondness, inasmuch as it requires (at least) two things standing apart from one another, thereby embodying Secondness as that which is in respect to some Other. The morphosyntactic diminution of conjunction in all contexts in Tamil is another Iconic representation of the negative constraint [-2].

In the paradigmatic realm Tamil, especially in more conservative dialects, like Jaffna Tamil in northern Sri Lanka, display the negative verb conjugations mentioned previously, which are a conspicuous typological trait of the Dravidian language family. In most spoken Tamil dialects on the Indian mainland, the negative conjugation is nowadays limited to the future habitual tense and to the literary language; but it is still significant enough typologically, historically, and dialectally to merit mention. Most languages indicate the negative by inclusion of some explicit negating word + the affirmative form, as in English *don't/doesn't* + verb, Spanish *no* + tensed form of verb, Russian *ne* + inflected form of the verb, Chinese *bu-* or *mei-* plus the verb, and so forth. This pattern typically extends to non-finite forms, like participles and verbal nouns, as with English *X-ing/not* + *X-ing*, *having X-ed/having* + *not* + *X-ed*, Spanish past participle/*no* + past participle (as in *hecho/no hecho*, ‘done/not done’), gerund/*no* + gerund (*haciendo/no haciendo*, ‘doing/not doing’), etc. The overall semiotic effect is a paradigmatic representation of negation as not-affirmation, or as a contrastive Other to the affirmative. Negation being a Secondness, as we have already seen, such paradigmatic pairings involving a strong (negative) and weak (affirmative) element exemplify the degenerate Secondness/[12] so intrinsic to Chinese grammar.

But we would expect Secondness in all its forms, including degenerate varieties thereof, to be minimized wherever the negative constraint [-2] is operative. As a result, Tamil has transformed negation into a separate system from affirmation, where the explicit paradigmatic pattern *A/not A* is not in evidence. There are, especially in formal registers, a significant variety of strategies for forming the simple negative finite tenses, whereas for the formation of their affirmative counterparts, only a single strategy exists.

In formal written Tamil, and, up until roughly a century ago, in the spoken language of educated people, a full negative conjugation consisting of the verb root + personal endings, bereft of the tense infix, sufficed for simple negation in past, present, and future. The contrast between this paradigm and the corresponding past,

present, and future tenses is shown on Table 11:

Table 11. Literary/formal Tamil simple negative and affirmative conjugations (cey-/ceyy-, ‘do, make’)

	Negative	Present	Past	Future
1st Sing.	<i>ceyy-ēn</i>	<i>cey-kir-ēn</i>	<i>cey-t-ēn</i>	<i>cey-v-ēn</i>
2nd Sing.	<i>ceyy-āy</i>	<i>cey-kir-āy</i>	<i>cey-t-āy</i>	<i>cey-v-āy</i>
3rd Sing. Masc.	<i>ceyy-ān</i>	<i>cey-kir-ān</i>	<i>cey-t-ān</i>	<i>cey-v-ān</i>
3rd Sing. Fem.	<i>ceyy-āḷ</i>	<i>cey-kir-āḷ</i>	<i>cey-t-āḷ</i>	<i>cey-v-āḷ</i>
3rd Sing. Formal	<i>ceyy-ār</i>	<i>cey-kir-ār</i>	<i>cey-t-ār</i>	<i>cey-v-ār</i>
3rd Sing. Irrational	<i>ceyy-ātu</i>	<i>cey-kir-ātu</i>	<i>cey-t-ātu</i>	<i>ceyy-um</i>
1st Plural	<i>ceyy-ōm</i>	<i>cey-kir-ōm</i>	<i>cey-t-ōm</i>	<i>cey-v-ōm</i>
2nd Plural	<i>ceyy-īrkaḷ</i>	<i>cey-kir-īrkaḷ</i>	<i>cey-t-īrkaḷ</i>	<i>cey-v-īrkaḷ</i>
3rd Plural Rational	<i>ceyy-ārkaḷ</i>	<i>cey-kir--ārkaḷ</i>	<i>cey-t-ārkaḷ</i>	<i>cey-v-ārkaḷ</i>
3rd Plural Irrational	<i>ceyy-ā</i>	<i>cey-kinr-ana</i>	<i>cey-t-ana</i>	<i>ceyy-um</i>

Note that, with this system, instead of the negative being represented by the affirmative plus some negating morpheme, word or words in the “*A/not A*” mold, the negative is simply the verb stem plus personal endings, bereft of overt tense marking, with no explicit negative marker of any kind and, therefore, no direct formal contrast with the affirmative.

In most dialects of modern Tamil, the formation of the negative has changed, such that only the future negative exhibits a full conjugational paradigm complete with tense and person/number affixes. It is formally entirely different from the future affirmative, as shown on Table 12.

Table 12. Modern Tamil affirmative and negative future (cey-/ceyy-, ‘do, make’)

	Future Affirmative	Future Negative
1st Sing.	<i>cey-v-ēn</i>	<i>ceyya-māṭṭ-ēn</i>
2nd Sing.	<i>cey-v-āy</i>	<i>ceyya-māṭṭ-āy</i>
3rd Sing. Masc.	<i>cey-v-ān</i>	<i>ceyya-māṭṭ-ān</i>
3rd Sing. Fem.	<i>cey-v-āḷ</i>	<i>ceyya-māṭṭ-āḷ</i>
3rd Sing. Formal	<i>cey-v-ār</i>	<i>ceyya-māṭṭ-ār</i>
3rd Sing. Irrational	<i>ceyy-um</i>	<i>ceyy-ātu</i>
1st Plural	<i>cey-v-ōm</i>	<i>ceyya-māṭṭ-ōm</i>
2nd Plural	<i>cey-v-īrkaḷ</i>	<i>ceyya-māṭṭ-īrkaḷ</i>

3rd Plural Rational	<i>cey-v-ārkaḷ</i>	<i>ceyya-māṭṭ-ārkaḷ</i>
3rd Plural Irrational	<i>ceyy-um</i>	<i>ceyy-ātu</i>

In this case, and in contrast to the original negative conjugation shown on Table 11, the future negative, except in the third person irrational, exhibits an infix, *-māṭṭ-*, which could be construed as a tense marker, on analogy with *-v-/-pp-* in the affirmative. Except for the third person irrational, the person/number affixes are identical. The stem for the future negative, however, is the infinitive, not the finite stem. Thus, although these two constructions are similar in some respects, they by no means conform to the morphosemiotic pattern *A/not A*.

As for the present and past tense negatives in the modern non-literary language, they are most usually formed by affixing *-illai*, the negative of *iru-*, ‘be,’ to the infinitive stem, with no tense or person/number affixation at all. Thus, e.g., *pōkavillai* (from *pōka*, ‘to go’) can mean ‘I don’t/didn’t go, you don’t/didn’t go, we don’t/didn’t go’, etc. This singular form for all persons and numbers in both the negative past and present stands in complete morphosemiotic contrast to the many forms in the affirmative paradigm, where both past and present, as well as a great variety of persons, numbers, honorifics, and cognitive qualities (rational vs. irrational, in the third person) are explicitly marked.

In addition, non-finite verb forms in Tamil exhibit morphology that does not conform to the *A/not-A* morphosemiotic pattern. The converbial or conjunctive participle mentioned previously in connection with conjoined clauses has both an affirmative or a negative form, which in English would be represented as *having A-ed/ having not A-ed*, again, in conformity with the *A/not-A* pattern. But in Tamil, whereas the affirmative conjunctive participle is formed using a variety of rules, depending on which of Tamil’s seven verb classes the verb belongs to (e.g., the stem + *-i* [Class 1] or the stem plus *-ttu* [Class 6]), the negative is formed across all verb classes by adding *-āmal* to the infinitive (e.g., from Class 6 stem *pār-*, ‘see’ [infinitive *pārkkā*], we have affirmative converbial *pār-ttu* and negative converbial *pārkk-āmal*). Much the same may be said of the Tamil affirmative and negative conditional forms, corresponding to the sense of ‘if Y does X’, which may be represented inflectionally by the affixing of *-āl* to the affirmative converbial stem for the affirmative conditional, and by the affixing of *-āviṭṭāl* to the infinitive stem for the negative conditional; thus from *cey-*, ‘make, do’ (converbial *cey-tu*, infinitive *ceyya*) *ceyt-āl*, ‘if Y does/makes’ versus *ceyy-*

āviṭṭāl, ‘if Y does not do/make’. Similar morphological asymmetries between the affirmative and negative are found with verbal nouns, verbal adjectives, and other non-finite forms.

In sum, the affirmative and negative verb paradigms, both finite and infinite, are represented in Tamil as completely different systems, showing none of the formal *A/not A* contrasts usual in Indo-European and many other language systems. We propose that this is yet another manifestation—paradigmatic rather than syntagmatic—of the systemic negative constraint [-2], because it serves to minimize formally the direct contrast between affirmative and negative.

Overall, Tamil morphosyntax exhibits pervasive evidence for the conditioning of both the positive constraint [+13] and the negative constraint [-2]. We would predict similar conditioning effects in other strongly agglutinating languages.

2.3.2 *Tamil lexicon*

Recalling that the positive constraint [+13] is most transparently manifest as variation on a single theme or idea, we consider whether this constraint is manifest in the lexicon per se. The most striking characteristic of the Tamil lexicon (a feature it has in common with many other South Asian languages, both Dravidian and Indo-Aryan) is its penchant for synonyms. Probably the most vexing aspect of many South Asian languages for the foreign learner is the need to become familiar with a significant number of exact synonyms for many objects. All languages have synonyms, of course, but more often than not, one word will be the standard usage, while the others will be relegated to literary, formal, idiosyncratic, or archaic contexts. The striking fact about Tamil is that many things have several synonyms associated with them that are exactly equivalent to each other in meaning and frequency of use, in addition to large numbers of additional synonyms that are distinguished for being more characteristic of either the literary or the colloquial register. One good example should suffice to illustrate this trait. The English word ‘body’ (meaning the physical body of a human or animal) has no synonyms in common use, like most common nouns in English. But its Tamil equivalents are legion, including, but not limited to: *akkam*, *ākāram*, *akkai*, *uṭampu*, *uṭal*, *sarīram*, *caṭakku*, *catiram*, *carīram*, *kaḷēparam*, *kaṭṭai*, *kāttiram*, *kaṭci*, *kāyam*, *kāyāpuri*, *malapāṇṭam*, *marukku*, *mēni*, *mūrtti*, *mūrttam*, *orru*, *pottai*, *poykkūṭu*, *tēkam*, *tēyam*, *uṭalam*, *uṭar*, and *vapu*.

Of these, the most common are *uṭampu*, *uṭal*, *carīram*, *kāyam*, and *tēkam*. Many

other words besides these may be employed in literary contexts, some of which are used only figuratively, or have reference to the body as, e.g., the receptacle of the soul; such usages include, but are not limited to, *muṭṭai*, *mōṭu*, *meyp̄poḷi*, *nataikkūtam*, *paṇṭam*, *paṭi*, *pukkil*, *tiḷam*, *uyirppoḷai*. And beyond these, there are long lists of synonyms for words with more specific, but related meanings, such as ‘dead body, corpse’ (*cavam*, *piṇam*, *pirētam*, *kuṇapam*, *mirutakam*, etc.) or ‘headless body’ (*aḷukuḷai*, *kavantam*, *matai*, *yūpam*, etc.).

This extraordinary proliferation of synonyms, some of which result from Sanskrit borrowings, and some of which are variants on a similar root resulting in distinct lexemes (*uṭal*, *uṭampu*, *uṭaram*, *uṭar*, or *mūr̄tti*, *mūr̄ttam*, e.g.), is evident throughout the Tamil lexicon, and evinces a compulsive love of synonyms that begs an explanation. Such sets of synonyms, inasmuch as they all represent the same Object, are variations on a single theme that we have come to recognize as the calling card of Firstness of Thirdness; and so we see that, not unexpectedly, the positive constraint [+13] is also operative at the level of the lexicon.

What about [-2]? The most interesting typological characteristic of the Tamil lexicon, besides the proliferation of synonyms, is the near-total absence of morphological adjectives and adverbs. All but a tiny handful of adjectives and adverbs in Tamil are derived from nouns, via the addition of suffixes *-āna* (adjectives) and *-āka* (adverbs; often reduced to *-ā* in the spoken language). The adjectives that are actually such as lexical entries include *nalla* (‘good’) and several color terms. This is in stark typological contrast to many other language families, including Indo-European; while all languages have derivational morphology whereby nouns and verbs may be transformed into adjectives and adverbs (such as *-ful*, *-ing*, *-al*, and *-ly*, among other, in English), adjectives and adjectives per se are typical word classes constitutive of many lexicons.

In Tamil (and the other Dravidian languages), however, “true” adjectives and adverbs are almost nonexistent, and such cannot be reconstructed in Proto-Dravidian, either.

That this typological quirk may be indicative of the negative constraint [-2] is suggested by the nature of adjectives and adverbs, and by what the presence of “true” morphological adjectives and adverbs in a lexical inventory might signify about the way that the language in question represents reality.

Adjectives and adverbs do not merely represent qualities; nouns and verbs

can both serve that purpose equally well. What adjectives and adverbs as lexical categories signify is a quality in relation to some Other, either a noun-Object or a verb-Object. In other words, they represent explicitly the notion that qualities must inhere in some Other. Therefore, in English, the fact that the adjective ‘red’ is primordial, and noun ‘redness’ and verb ‘reddden’ are both derived from it, implies that, for English speakers, the relational property of ‘red’ in inhering in some Other is taken as fundamental, prior to the notion of ‘red’ both as a Subject and as a Predicate. Moreover, the fact that English has many hundreds of morphological adjectives and adverbs in its lexicon implies that this notion is very widespread; and the same could be said of the large number of other languages for which this is the case.

But in Tamil, the reverse is true. Most words denoting qualities are morphological nouns (which may also play the role of predicates, as subject complements). Adjectives and adverbs are derived from them, suggesting iconically that the relational property of adjectives, as representing a quality inhering in some Other, is strictly subordinate and not semiotically prominent. This is tantamount to saying that Otherness, or Secondness, is not semiotically prominent, at least where adjectival and adverbial morphology is concerned, or that the constraint [-2] would indeed appear to be operative.

2.3.3 *Tamil phonology*

The major typological trait of Tamil phonology is the extreme paucity of obstruents, whereof the total number of phonemic stops and affricates is only six. Tamil has only one velar stop /k/, one dental stop /t̪/, one labial stop /p/, and one palatal stop /tɕ/, each of which exhibits at least four different allophones, depending on whether they occur in the initial, geminate-initial, intervocalic, or post-nasal position. /k/ exhibits at least six different allophones in most dialects, including three, [x], [ɣ], and [∅], in free variation intervocalically. Labial /p/ exhibits five different allophones, including both [β] and [w] intervocalically, and palatal /tɕ/ has five allophones, of which two, [tɕ] and [s], both occur initially. Two additional stops, a retroflex and alveolar, while not occurring initially, each exhibit at least three allophones corresponding to the other three phonetic environments. Rather unusual typologically is that neither voice nor aspiration are phonemically contrastive in Tamil, as also is supposed to have been the case in Proto-Dravidian; this is in stark contrast to Indo-Aryan languages, where both voice and aspiration are distinctive articulatory features.

What this inventory of Tamil stops and affricates shows is an unusually small inventory of phonemes coupled with a diverse array of allophones apiece. In English and most European languages, which typically boast large and diverse arrays of phonemic stops and affricates, the number of allophones per phoneme is usually fewer, with English /t/ and /d/ exhibiting, on average, 4 and 3 allophones, respectively, and Spanish /t/ and /d/ roughly 3 apiece. European languages also allow for much more diverse and complex consonant clusters than does Tamil, increasing the number of possible environments for allophonic variation to occur; yet Tamil still exhibits more variation, on average. This is another striking instance of variation on a theme that we have associated with the positive constraint [+13] in Tamil, this time operative at the level of phonology.

As for the negative constraint [-2], note that the paucity of phonemic stops and fricatives in Tamil, and in particular, the absence of [+/-voice] and [+/-aspiration] featural oppositions, means that Tamil lacks the most fundamental contrastive element found in consonant inventories: the fortis/lenis opposition. These terms have reference to energy levels in otherwise identical or nearly-identical pairs of consonant phonemes, energy that may depend upon the presence or absence of not only voice and aspiration but also on characteristics like glottalization and even length. In English, for example, the pairs /t/ - /d/ and /p/ - /b/ epitomize the fortis – lenis opposition; in initial position, for example, /t/ and /p/ are aspirated, and therefore significantly more energetic than their voiced equivalents. The fortis-lenis opposition is widespread, and is found in Indo-Aryan, Munda and Tibeto-Burman languages elsewhere on the South Asian Subcontinent. It is also characteristic of Chinese phonology. Such an opposition is a paradigmatic instantiation of an unequal dichotomy, a pairing we recognize to be Peirce's degenerate Secondness, or [12]. [12] being a form of Secondness, we may represent its absence, in the case of Tamil's nonexistent fortis-lenis opposition, as [-2] (although in more precise terms, [-12] would be more accurate).

We have seen with Tamil strong and consistent data suggesting the presence of [+13] and its negative concomitant [-2] as Categorical constraints at every level of grammar that we have examined, and in both the syntagmatic and paradigmatic subdomains. In general, the characteristic of variation on a single theme or idea, the cardinal manifestation of [13], is responsible for all of the typologically prominent features of Tamil morphosyntax, lexicon, and phonology. We summarize the manifestation of the semiotic structures resulting from the conditioning constraints

[+13] and [-2] as follows:

[+13]:

Morphosyntax:

- Enormous variety of forms associated with both nouns and verbs as a result of agglutinative morphology

Lexicon:

- Exceptionally large numbers of synonyms

Phonology:

- Very small number of phonemic obstruents coupled with large numbers of allophones

[-2]:

Morphosyntax:

- Absence of periphrasis
- Nouns and pronouns conjoined by multiplicative postclitic -um
- Clauses conjoined by converbial participles rather than an explicit conjunction
- Absence of formal *A/not A* contrast between affirmative and negative verbs

Lexicon:

- Near-total absence of morphological adjectives and adverbs

Phonology:

- Absence of fortis-lenis opposition

2.4 Vedic Sanskrit

The other classical language of South Asia, Sanskrit (whose name means “refined”, as opposed to “Prakrit” [a blanket term for the many literary dialects, including Pāli, derived from Sanskrit, from which the Middle Indic “Apabhramas”—the predecessors of modern Indo-Aryan languages—were descended], which means “unrefined, common”), is not only the ultimate wellspring of all of the modern Indo-Aryan languages with their hundreds of millions of speakers, it is also, alongside Hittite, the oldest attested Indo-European language, and by far the best-documented of all the primordial Indo-European languages. Although Sanskrit is still spoken today by a few thousand Indian Brahminical scholars, it is technically a dead language in the same sense of Medieval Latin during the Middle Ages; the last authentic form of

the language was so-called “Buddhist Hybrid Sanskrit”, which was preceded by the Classical Sanskrit of Kālidāsa, the epic Sanskrit of Valmiki, and the Vedic Sanskrit of Angiras and the other mysterious sages who composed the Vedic hymns and other early Sanskrit literature, beginning sometime in the second millennium BC. Vedic Sanskrit being the most conservative form of this language, and withal the version least apt to display influences from foreign idioms, especially Dravidian, we will try to determine what Categorical constraint may be responsible for the semiotic structures evident in Vedic Sanskrit.

Vedic Sanskrit being a dead language, and one whose available sampling, though very large, is nonetheless skewed in favor of stylistic preferences found in votive poetry, our analysis is doubtless incomplete, based, as it must be, on incomplete data. Nevertheless, thanks to the extraordinary accuracy with which every detail of Sanskrit grammar has been preserved, not to mention the copious amount of literature that has come down to us, we probably have more semiotic data for Sanskrit, including Vedic Sanskrit, than for any other ancient language, save possibly Hebrew.

2.4.1 Vedic morphosyntax

That said, what can be posited as a Categorical constraint for Vedic Sanskrit? The most prominent typological trait of Sanskrit morphosyntax is the almost unbelievable luxuriance of its declensional and conjugational paradigms, which enables a system of agreement unparalleled in any other language. Not only subject-verb agreement, but also adjective-noun agreement constitute the morphosyntactic heart and soul of Sanskrit, and verbs, nouns, pronouns, and adjectives are all divided into large numbers of classes with very distinctive conjugational and declensional patterns. Nouns, for example, are broadly divided into various vowel stem (using so-called “thematic vowels” as intermediaries between stem and affix) and consonant stem classes, which in turn are organized into a large number of different declensions (far more than the five familiar to students of Latin!) displaying eight cases across three numbers, singular, dual, and plural. Pronouns are no less ornate than nouns, exhibiting all of the same cases and numbers. Adjectives agree in all particulars with nouns, and include an unusually diverse array of demonstrative adjectives. Verbs, meanwhile, are conventionally divisible into ten classes, some of which involve a thematic vowel in finite forms, others a root stem, and others various infixes (which would appear to be analogous to thematic vowels). Morphologically, Sanskrit verbs

are said to present four tense systems based on four different forms of the root used in finite conjugation, viz., the present (including the present, imperfect, imperative, and optative), the perfect, the aorist, and the future (including the future per se as well as the conditional). All of these verb forms typically distinguish three persons and three numbers via subject agreement morphology, for a total of nine forms per tense. Not only that, verbs typically have both an active (*parasmaipada*) and middle (*ātmanepada*) inflectional voice, in addition to a passive.

Because of such reliance on agreement, signaled by such elaborate inflectional morphology, word order in Sanskrit is comparatively free; with practice, the student learns to associate nouns, verbs, and adjectives via inflectional agreement alone.

This state of affairs would seem to suggest a tentative hypothesis, that Vedic Sanskrit is constrained primarily by [+3], since agreement, as we have elsewhere noted, is the strategy of meaning association emblematic of Thirdness. If [+3] is indeed the positive constraint conditioning Sanskrit, then what other characteristics might we expect? Thirdness, as earlier set forth, is the Category that, in the broadest sense, embraces all phenomena having to do with standing as intermediary between a First and a Second, so as to establish a relationship between them. Symbols (as opposed to Icons and Indexes) are the Signs that best embody Thirdness. The number three (i.e., the trichotomy), is one of the calling cards of Thirdness (in contrast to the oppositional dichotomies of Secondness and the unitarity of the First). For where there is a Third, there must perforce be a First and a Second also. Representation (Peirce's term of choice for the intermediacy of Thirdness), in particular representation via Symbols, Legisigns, and Arguments, is Thirdness par excellence. Other secondary phenomena that arise from Thirdness include the mediation, habit, law, plurality, evolution, continuity, and regularity mentioned previously, as well as final causation, thought, and growth. Any of these may be indicative of Thirdness as a constraint.

The agreement that perfuses Sanskrit morphosyntax is a manifestation of [+3] in the syntagmatic subdomain. What of the paradigmatic? We note that Sanskrit paradigms are organized, conspicuously and consistently, into trichotomies. The most fundamental of these, perhaps, is the trichotomy of number (singular, dual, plural), which three entries correspond very neatly to Firstness, Secondness, and Thirdness, respectively. This trichotomy applies to nouns, verbs, and adjectives, and is assumed to have also been a morphosyntactic characteristic of Proto-Indo-European, although the dual number has survived only in a few other attested Indo-European languages

(in ancient Greek and Gothic, for example; nowadays, it persists in a few conservative Indo-European languages, like Lithuanian and Slovenian). Considering verbs alone, we find many more trichotomies, including the first/second/third person (which are strictly marked as such, i.e., three and only three forms apiece in the singular, dual, and plural¹¹) and the active/middle/passive voice. Nouns are classified as masculine/feminine/neuter.

Because Sanskrit verb paradigms are assembled according to intuitions about tense, mood, aspect, and other verbal properties acknowledged by Western (and, *mutatis mutandis*, Indian) grammarians, the significance of subtler trichotomies—i.e., those that do not square with such purely grammatical features—is apt to be overlooked. A very good example of a morphosemiotic trichotomy in Vedic Sanskrit that has passed unnoticed by linguists is the very singular manner in which the external contours of the verb stem may be represented (i.e., relative to other affixal morphology). Broadly speaking, the verb stem may take three forms relative to the affixal morphology that follows it: it may appear as the stem itself (including various gradational forms of the internal vowel, which we shall discuss separately further on), it may receive the so-called “augment” *a-*, and it may be reduplicated. Thus, e.g., root *kṛ-*, ‘make, do’, may occur as a bare stem (*kṛ-*, *kar-*, or *kār-*), with augment *a-* (*akṛ-*, *akar-*, or *akār-*), or reduplicated with *ca-* (*cakṛ-*, *cakar-*, or *cakār-*; Sanskrit root reduplication operates according to rather complex rules; in this case, palatal [c] normally substitutes for reduplicated [k]). This constitutes yet another formal trichotomy, but what is its purport, if anything? Because these three forms of the verb root do not correspond neatly to single tenses, moods, etc., but instead occur in various grammatical contexts, it is tempting to assume that they have no specific grammatical meaning. But we are concerned with semiotic and not grammatical structures; under our assumption that formal differences signpost semiotic distinctions, we are prompted to look for semiotic structuring undergirding this formal trichotomy.

Reduplication is overwhelmingly associated with the perfect system in Vedic Sanskrit. It is important to note that the Vedic language in its pristinity was organized as a set of systems (of which the “perfect” was one), each of which encompassed subsystems marking morphological voice and mood, and sometimes tense/aspect. The perfect system encompasses both the *parasmaipada*/active and *ātmanepada*/middle voices, as well as the indicative, subjunctive, injunctive, optative, and imperative moods. It was not, however, a “tense” in the sense understood by speakers of modern

Indo-European and many other language families. According to Kiparsky (1998, p. 4), the perfect could denote events either in the present or past time.¹² It denoted what Kiparsky calls “stative presents” applicable to “achievement verbs” like *ciketa* (‘knows’, *cit-*), *jujoṣa* (‘enjoys’, *juṣ-*) *cakāna* (‘likes’, *kan-*), *bibhāya* (‘fears’, *bhī-*), *dadhāra* (‘holds’, *dhr-*) and *babhūva* (‘is’, *bhū-*). The contrast between such uses of the perfect and the simple present may be seen in utterances where the two co-occur in the same complex predicate, as in

ka īṣate *tujyate* *ko bibhāya*
 who flee-3rd Sing PRES rush-3rd Sing PRES who fear-3rd Sing PERF
 ‘Who is fleeing and rushing, who is afraid?’

Note that the first two verbs, ‘flee’ and ‘rush’, denote dynamic, transient states, whereas the third, ‘fear’, denotes a more or less permanent condition, and withal one characterized by some state of consciousness. Much the same could be said of most of the other “achievement verbs” listed previously, including *bhū-* (meaning “being” in the sense of possessing some habit, characteristic, or quality, as well as “becoming”, in contrast to the root *as-*, which is the Sanskrit copula).

As for the past time contexts of the Vedic perfect, Kiparsky notes that “the most salient perfect-specific function is to introduce an existential or universal quantification over past times. Predicates interpreted distributively, denoting multiple events, nearly always have the perfect in reference to past time” (Kiparsky, 1998, p. 4). As an example showing the contrast of the one-event imperfect and the multiple-event perfect, Kiparsky cites the following Vedic passage:

dame-dame ... agnir hotā ni sasādā *yajīyān.* *agnir hotā*
 in every house Agni priest down sit-3rd Sing PERF more sacrificing Agni priest
ni asīdad *yajīyān* *upasthe mātuh*
 down sit-3rd Sing IMPERF more sacrificing lap-Loc mother-Gen
 ‘Agni the expert priest sat down [multiple events] in every house. Agni the expert priest sat down [single event] on his mother’s lap.’

Kiparsky also observes:

The distributive reading [for the perfect] typically occurs with plural or collective subjects or objects. Indeed, with universally quantified plurals, the perfect is mandatory.... A telling contrast occurs in the creation hymn 10.129. It first says that certain things did not exist—being did not exist, death did not exist (*āsit*, imperfect) – and then concludes: nothing existed (*āsa*, perfect).

The perfect is obligatory, even with singular nominal arguments, when the verb is modified by an adverb of quantification, either universal (such as “always” ...) or existential, such as “many times”.... The perfect is also the normal past tense of generic and habitual sentences.... In contrast, questions about specific past occasions, asking for particular answers... have imperfect tense. (Kiparsky, 1998, pp. 5-6)

All of which is to say that the perfect in its past time usage is concerned with plurality, multiplicity, and habit. These, together with the notions of permanence and conscious intent involved in present time usages, suggest very strongly that the perfect—and, consequently, the reduplication emblematic thereof—is in fact representative of Thirdness.

The second morphological form of the stem, with augment a-, is characteristic of the other two past tenses in Vedic, the imperfect (as in *asīdad*, ‘sat’, noted above) and the aorist. The use of the imperfect is to represent past events relative to the present (the imperfect constituting a part of the “present system” in Vedic). The aorist, by contrast, is somewhat more multifarious. Kiparsky notes four different uses of the aorist, viz., the “perfective aspect” (i.e., the completive sense), the “immediate past”, the “relative anteriority of a subordinate clause with respect to the main clause”, and “a statement of fact”. Remarking on the seeming disparity of these uses, Kiparsky observes:

The link between these four functions is certainly not obvious. What exactly do the temporal meanings “recent past” and “relative anteriority” have to do with one another? What does either of them have to do with the discourse function of “statement of fact”? And what do any of these have to do with the telicity [i.e., perfective/completive] that the aorist marks in non-indicative contexts? Part of the problem is to discover what *kind* of connection we should be looking for: a basic meaning from which the several functions are synchronically derivable? Or a natural diachronic path that connects them, grounded in principles of language change? And the answer to that depends on whether the tenses have a single lexical meaning with structural ambiguity, a single meaning with different

pragmatic uses, or are genuinely polysemous. (Kiparsky, 1998, p. 3)

Part of the reason for the perplexity is the insistence on seeing the aorist as a purely grammatical entity instead of a more generally semiotic one. In Peircean semiotic terms, there is a very simple unifying characteristic, not only of the seemingly divergent uses of the aorist, but also of the imperfect. Both of these are “past tenses”, with the imperfect specifically denoting single events in the past (in contrast to the multiplicity or habituality of the perfect) and the aorist denoting both immediate and anterior past. Additionally, the aorist may connote completion (perfectivity, a confusing term in light of the fact that the perfect, per se, is not in fact perfective in Vedic!) as well as factuality. All of these characteristics—the past (especially in the immediate or instantiated sense), completion, and fact—are all prominent manifestations of Secondness, and from this we may confidently conclude that the augment a- that signalizes both the imperfect and the aorist is in fact a morphosemiotic signpost for Secondness.

The bare root form, unaugmented and unreduplicated, is characteristic of the simple present tense in Sanskrit as well as a number of other tenses lacking either a past or completive sense, or any notion of multiplicity or habituality, such as the present optative, and the injunctive. This last, virtually non-existent in post-Vedic Sanskrit, according to Kiparsky, “can assume virtually any temporal or modal value in context ... It is clear that the injunctive should be treated as a form which has no tense and mood specifications”. Both the notion of the present, in general, and the absolute absence of tense and mood are suggestive of a semiotic emphasis on the mere quality of meaning of the root itself, or in other words, of Firstness.

Thus the trichotomy of root forms—root-augmented/root-reduplicated/root—present yet another a tidy representation of the Categories embedded in Vedic grammar, and withal a trichotomy that does not square with the conventional paradigms laid out by grammarians. The semiotic structure of the trichotomy/[+3] does not, in this case, coincide with the grammatical structure, but it is evident nonetheless.

2.4.2 *Vedic lexicon*

At the lexical level, the most conspicuous typological characteristic of Sanskrit (as with many other Indo-European languages) is the tripartite structure of most “big”

words (i.e., words that may serve as Subjects and Predicates: nouns, verbs, and adjectives). This structure, familiar to all comparative Indo-Europeanists, consists of a stem (derived from an abstracted root) + thematic vowel/[word] infix + inflectional affix. In the case of nouns and adjectives, the thematic vowel may be *-a-*, *-ā-*, *-i-*, *-ī-*, *-u-*, or *-ū-*, and is followed by a set of endings that signify case and number. Unlike Tamil and many agglutinative languages, there is no discrete morpheme to indicate number. In the case of verbs, only Class 1 utilizes a thematic vowel (*-a-*); other classes are distinguished by a range of infixes, including *-na-*, *-no-*, and *-aya-* interposed between stem and inflectional affixes marking person and number. Some examples of these forms are shown on Table 13:

Table 13. Examples of stem-thematic vowel/infix-affix structure of Sanskrit nouns, adjectives, and verbs

a-stem noun (nominative and accusative singular)

dev-a-ḥ

god-[thematic *a*]-Nominative Singular

dev-a-m

god-[thematic *a*]-Accusative Singular

i-stem adjective

śuc-i-ḥ

clean-[thematic *i*]-Nominative Singular

śuc-i-m

clean-[thematic *i*]-Accusative Singular

u-stem adjective

mṛd-u-ḥ

soft-[thematic *u*]-Nominative Singular

mṛd-u-m

soft-[thematic *u*]-Accusative Singular

Class I verb (thematic a)

nay-a-si

lead-[thematic *a*]-2nd Singular ('you lead')

nay-a-nti

lead-[thematic *a*]-3rd Plural ('they lead')

Class V verb (-nu-/-no- infix)

su-no-si

press-[infix -no-]-2nd Singular ('you press')

su-nu-taḥ

press-[infix -nu-]-3rd Dual ('you two press')

Class VIII verb (-u-/-o- infix)

kar-o-mi

do-[infix -o-]-1st Singular ('I do')

kur-u-tha

do-[infix -u-]-2nd Plural ('you do')

Class IX verb (-nā- infix)

krī-ṇā-ti

buy-[infix -nā-]-3rd Singular ('(s)he buys')

krī-ṇa-nti

buy--[infix -nā-]-3rd Plural ('they buy')

Not all Vedic nouns and verbs involve a thematic vowel or infix; some consist of a stem + affix. And many of those that do—especially among the verbs—do not always present entirely consistent, “neat” results, owing to the rules of sandhi, evolutionary accident, and other factors. This, as we have elsewhere noted, is a hallmark of semiotic systems and of the randomness and spontaneity associated with the element of Firstness always contained therein (exceptionless regularity being a contrivance of reductionism only, not a feature of the existent semiotic universe). Notwithstanding, a majority of Sanskrit nouns, adjectives, and verbs probably do conform to this tripartite template, at least among unmarked forms like the simple present. Of Sanskrit’s 10 verb classes, for example, only two are truly athematic, i.e., lack any kind of thematic vowel or epenthetic infix. Among nouns and adjectives, thematic *a*-stem nouns are by

far the most numerous class, with *i*-stem and *u*-stem (along with their mostly feminine counterparts, *ā*-, *ī*-, and *ū*-stem) making up a very substantial portion of the rest. A few athematic nouns, like *nṛ*-, ‘man, person’, and *hṛd*-, ‘heart’, also had thematic forms *nar-a*- and *hṛd-a(ya)*, respectively, which became the standard forms by the classical stage of the language. The canonical form stem + thematic vowel/infix + inflectional affix thus is the dominant lexical template even in the Vedic phase of Sanskrit.

Nominal and verbal stems in Sanskrit (as in many other Indo-European languages, from ancient Greek to modern Spanish) are not words themselves, since they represent, not independent, free-standing lexemes (even in potential), but instead ideas of Objects and Qualities which may only be fully instantiated lexically by combining with affixes.

Something like this phenomenon is found in some other languages and language groups; in Tamil and other Dravidian languages, for example, a distinction is made by grammarians between the so-called “direct” and “oblique” forms of nouns in the singular, such that the “bare” form of the noun—i.e., the word itself, unadorned by any affixal material—is used for the nominative or “direct” case, while a separate “oblique” form featuring a slightly modified final syllable is the form to which the non-nominative or “oblique” case endings are added. However, this oblique form is limited to the singular; in the plural, the plural marker is added to the direct stem (sometimes modified per the requirements of sandhi). Moreover, the “bare” oblique stem in the singular may serve as the genitive case. In contrast to such instances from non-Indo-European agglutinative declensional systems, the Sanskrit root almost never can appear unadorned or unmodified; it is a true abstraction, signifying iconically a quality which can only be reified by association with other morphology.¹³ On the other hand, the inflectional affix—case/number for nouns, person/number for verbs—situates the root, and its denotatum, with respect with some Other, be it a grammatical subject (verbs), some Predicate (nominative, accusative, and dative cases), some other argument of the Predicate (instrumental, ablative, locative), etc. Thus the root/stem and the inflectional affix represent Firstness/[1] and Secondness/[2], respectively. They are brought into association by a third overt morphological element, the thematic vowel or infix, whose only discernible purpose is to stand as intermediary between stem and inflection, a syntagmatic Icon of Thirdness/[3] at the lexical level. Overall, then, most nouns, verbs, and adjectives are perfect iconic representations of the trichotomy First/Second/Third characteristic of Symbols and Thirdness generally

(recall that [3] always involves both [2] and [1]). Just as Chinese words tend to embody the unequal dichotomy of [12], Sora words the blurring of lexical boundaries and tendency for absolute unity of [1], and Tamil words the luxuriant variation on a theme characteristic of [13], so Sanskrit words embody the trichotomy Quality-Reaction-Relation typical of [3], and thus many Sanskrit lexemes are formally conditioned by [+3], as predicted.

At the lexical level, the positive constraint [+3] is also manifest on the paradigmatic axis. Consider the two ablaut grades in Sanskrit, so-called *guṇa* and *vṛddhi*, which, in combination with the zero grade, constitute a trichotomy of form whereby root forms (the zero grade) may be transformed into common nouns (*guṇa*), as well as adjectives and nouns derived from such adjectives (*vṛddhi*).

Table 14 shows some examples of the Sanskrit system of ablaut, which is presumed to be reflective of a primordial Indo-European system.

Table 14. Sanskrit zero-grade, *guṇa*, and *vṛddhi*

Zero-grade	Guṇa	Vṛddhi
<i>div</i>	<i>dev</i>	<i>daiv</i>
<i>śuc</i>	<i>śoc</i>	<i>śauc</i>
<i>vid</i>	<i>ved</i>	<i>vaid</i>
<i>diś</i>	<i>deś</i>	<i>daiś</i>
<i>bhṛ</i>	<i>bhar</i>	<i>bhār</i>

These three grades appear in various contexts, and while from a strictly comparative historical perspective, it is in fact the “*guṇa*” grade that is the base form, Sanskrit has reconditioned them to represent *guṇa* and *vṛddhi* as additive, whereby the first is represented formally either by vowel lengthening or lowering, and *vṛddhi* by lengthening or diphthongizing the *guṇa*. In general, the pattern observed by the trichotomy *div-/dev-/daiv-* is both instructive and typical. The zero grade represents the verb root form, the general notion associated with the predicate (in this case, being bright or shining), as well as nouns more or less intrinsic to the meaning, while the *guṇa* grade is deployed in concrete nouns denoting discrete objects that directly embody this quality, and the *vṛddhi* grade is used for higher-order nouns and adjectives which often consciously or deliberately embody or employ the characteristics of the denotatum of the *guṇa* grade. The zero-grade and *guṇa* grade forms are also deployed for verb formation. Thus *div-* is not only the root form (and

conjugational stem for) the predicate ‘be bright, shine’ (as well as a range of other meanings, including ‘wager’, ‘spread’, ‘praise’, and ‘trifle’), it also is constitutive of the noun *diva-*, ‘day, heaven, sky’ (i.e., the very embodiment of the state of being bright or shining, in general). In finite verb stem forms, it is mostly limited to the simple present tense. *dev-*, the guṇa grade, is used not only in concrete nouns denoting tangible objects, like *deva-*, ‘god, divine being, ruler’, it is also characteristic of the stems of verbal nouns like the infinitive (*devitum*), tenses involving past or completive senses like the aorist (*adevit*) and the perfect (*dideva*), and more complex, marked usages like the desiderative (*didevishati*), the causative (*devayati*), and the conditional (*adevishyat*). The vṛddhi grade denotes characteristics or qualities of objects denoted by guṇa-grade words, and sometimes also nominalizations of such characteristics. Thus *daiva-* as an attributive means ‘sacred, heavenly, belonging to or sacred to the gods’ (all attributes of *deva-*), it may also, as a noun, signify ‘divine power’, ‘destiny’, and ‘religious offering’, among other things.

Because of the immense complexities of Sanskrit verbs, and the fact that different tenses and modalities are—as we have already seen—typically signaled by a set of markers, including (in addition to root vowel grade) reduplication, augmentative prefix *a-*, and various infixes like *-n-* and *-s-*, it is nonetheless possible to generalize that the guṇa grade almost always is used in the formation of infinitives and the causative, and frequently (though by no means universally) in the aorist. The root *diś-*, ‘show, point out, exhibit’, for example, has infinitive *deṣtum* and causative *deśayati*, as expected—but the aorist (a so-called “sigmatic aorist” because of the sibilant infix) includes forms like *adikṣat* and *adikṣi*, where the aorist is denoted not by root vowel gradation but by the prefix *a-* plus the sigmatic infix *-s-* (the root *diś-* here permutes to *dik-*, one of its alternate forms). The present tense, meanwhile, has two different forms; one of them, found only in the R̥gveda, is athematic and involves reduplication and the guṇa grade (*dideṣṭi*), while the other is thematic and involves the zero grade (*diśati*). This latter form is the only one found in post-Vedic Sanskrit.

Outside the messiness of verb conjugation, however, the various nominalizations of *diś-* follow more predictable patterns of vowel gradation meaning association, viz., *diś-* (also realized as *dik-* and *dig-*), ‘direction, cardinal point, region’ (general meaning, derived directly from meaning of predicate root), *deśa-*, ‘spot, place, region, country, kingdom’ (concrete object, the embodiment of the base meaning), *daiśika-*, ‘relating to space or to a country; local, provincial, national; native’ (attributive

embodying the properties of *desa-*).

The root *vid-*, ‘know, perceive, understand’, is manifest in the zero grade as *vid*, ‘knowing, understanding’ and as *vida-*, ‘knowledge, discovery’. In the guṇa grade, *veda-* means ‘knowledge, lore, sacred knowledge’ in the sense of some tangible compilation or repository (as in the *ṛg-veda*, literally, ‘verse-knowledge’). The vṛddhi grade form *vaida-*, meanwhile, means, in the attributive sense, ‘learned, knowledgeable, having to do with a wise man’, while the noun means, simply enough, ‘wise man’. As a verb, zero grade *vid-* occurs in the present tense (*vidati*), where it alternates with the guṇa grade (*vedati*), and is also used, in conjunction with augment *a-* and stem reduplication, for the aorist (*avividat*). The guṇa grade, as usual, is deployed in the infinitive (*veditum* or *vettum*) and in the causative (*vedayate*).

Still another example, *śuc-*, illustrates a separate zero grade-guṇa-vṛddhi series, *u/o/au*. The root *śuc-* has several meanings, including ‘suffer, grieve’, ‘shine’, and ‘purify’. From this root arise zero-grade nouns like *śuc-*, ‘flame, glow; brightness; pain; purity’ and *śuci-*, ‘honesty; purification’; both also have adjectival meaning ‘shining’, and *śuci* also means ‘holy, pure’. Guṇa forms include *śocana-*, ‘grief, sorrow’, *śoci-*, ‘flame, glow’, and *śocis*, ‘flame, glow, color, splendor, beauty’. Vṛddhi form *śauca-* means ‘cleanness, purity, purification; integrity, honesty’, and is also found in combined forms like *śaucācāra*, ‘purificatory rite’ and *śaucakalpa*, ‘mode of purification’.

The use of these grades in verb forms of *śuc-* is, as always, complex, and sometimes depends on which meaning of the root is intended. Thus for the meaning ‘suffer, grieve’, the present tense is usually guṇa (*śocati*), whereas for ‘be bright, be pure’, the present is zero grade (*śucyati*). The aorist is zero grade (*aśucat*), whereas the perfect is guṇa (*suśoca*). As always, the infinitive (*śoktum* or *śocitum*) and the causative (*śocayati*) are guṇa.

Some of the zero grade, guṇa, and vṛddhi uses of *div-*, *diś-*, *vid-*, and *śuc-* are summarized on Table 15.

Table 15. Zero grade, guṇa, and vṛddhi uses of *div-*, *diś-*, *vid-*, and *śuc-*

Root	Zero Grade	Guṇa	Vṛddhi
<i>div-</i> , ‘be bright, shine’	<i>diva-</i> , ‘day, heaven, sky’	<i>deva-</i> , ‘god, divine being’	<i>daiva-</i> , ‘sacred’
<i>diś-</i> , ‘show, point out’	<i>diś-</i> , ‘direction, cardinal	<i>deśa-</i> , ‘spot, place,	<i>daiśika-</i> ,

	point'	country, region'	'local, national, native, provincial'
<i>vid-</i> , 'know, perceive, understand'	<i>vida-</i> , 'knowledge, discovery'	<i>veda-</i> 'knowledge, lore, sacred knowledge'	<i>vaida-</i> , 'learned; wise man'
<i>śuc-</i> , 'grieve, shine, 'purify'	<i>śuc-</i> , 'flame, glow; brightness; pain; purity'; <i>śuci-</i> , 'honesty; purification; shining; pure; holy'	<i>śoci-</i> , 'flame, glow'; <i>śocana-</i> , 'grief, sorrow'; <i>śocis</i> , 'flame, glow, color, splendor, beauty'	<i>śauca-</i> , 'cleanness, purity, honesty'

The zero grade, denoting not only the abstracted root form, but also nouns and adjectives whose meaning is simplest and most proximate to the general root meaning, is everywhere suggestive of Quality or Firstness. This notion is given added credence by the fact that is most characteristically used for finite verb conjugation in association with the present tense, the tense associated with Firstness ("On Phenomenology", Peirce, 1998, pp. 149-150).

The guṇa grade is usually associated with more tangible instantiations of the root meaning, usually some bounded or measurable object, or a quality with more specifically quantifiable or measurable characteristics. Thus, for example, while *vida-* means 'knowledge' (in the general sense of the term), *veda-* means 'repository or body of knowledge'. The two most characteristic usages of the guṇa grade in verb formation are infinitives and the causative; the former, a verbal noun in Sanskrit which—like all nouns—is strongly indexical, and the latter, a version of the verb whereby the action is enjoined by one Subject upon another, or, in other words, where the Predicate is infused with the notion of Otherness or Reaction. Indexicality, Otherness, and Reaction all being aspects of Secondness, it seems clear that the guṇa grade, in its most general sense, is a sign of [2].

As for the vṛddhi grade, most vṛddhi nouns have the sense of "someone or something that has acquired the characteristic of the guṇa grade", while vṛddhi adjectives have the sense of "some complex, secondary trait that is characterized by the guṇa grade entry". Thus *daiva-*, as we have seen, means 'sacred, heavenly, belonging to or sacred to the gods' (adjectival) and 'divine power', 'destiny', and

‘religious offering’ (substantival). But notions like sacredness, belonging to the gods, religious offerings, etc., all carry such meanings purely by virtue of being assigned them by a speaker community, and by being activities created by cognitive acts. They are alike religious symbols, their sacred nature (as with most such phenomena) conferred by habit of association, and not necessarily as a consequence of any inherent sanctity. Again, *daiśika-*, the vṛddhi grade of *diś-*, invokes a set of properties (being local, or provincial, national, or native) that all arise from conscious deliberation. Mere ‘places’, ‘directions’, or ‘locations’ are all physical concepts that require no deliberation for their identity to be established as such. A mountain, a river, a lake, a field, etc., are purely natural phenomena, and a *deśa-* is a territory characterized by the same. But *daiśika-* refers to a person who bears the characteristics of belonging, in some sense, to a location or place, including especially polities devised by human thought. Here, too, the notions of habit, cognition, and deliberate human contrivance are paramount. So too with the vṛddhi grade *vaida-*, which endows human beings with the characteristics associated with *vid-/ved-*. In vṛddhi grade *śauc-*, too, these general notions are evident, inasmuch as ‘purity’, ‘purification’, ‘integrity’, and the like are all traits of human character, defined by moral values dictated by collective deliberation and long-established habits and norms of behavior. All of these characteristics—habit, thought, deliberation, and the like—belong to the ambit of Thirdness. The vṛddhi grade, therefore, is a paradigmatic representation of Thirdness, or in other words, the conditioning effects of the positive constraint [+3], at the lexical level.¹⁴

2.4.3 Vedic phonology

At the phonological level, Sanskrit consonants are not only phonemically diverse but also readily form clusters, such that Sanskrit words and texts (like English words) are easily recognizable in writing without vowels. Indeed, this intuition is encoded in all known South Asian writing systems with which Sanskrit, including Vedic, may be written, from primordial scripts like Brāhmi to modern scripts from Devanāgarī to Telugu to Sinhalese: all such scripts are abugida or alphasyllabic scripts, for which consonants are graphemically dominant, vowels normally being represented by subordinate graphemes, unless occurring word-initially (although modern Hindi has diluted this rule considerably). Many of these scripts (especially Devanāgarī, Malayalam, and Telugu) have elaborate rules for forming many distinctive compound graphemes to represent consonant clusters.

The Sanskrit consonant inventory includes 16 plosives, 5 nasals, two liquids, and nine fricatives and affricates. Admissible consonant clusters (including clusters crossing syllable boundaries) include *-str-*, *-sm-*, *-sp^h-*, *-shr-*, *-shtr-*, *-tsn-*, *-tpr-*, *-db^hr-*, *-ntr-*, *-nd^hr-*, *-mb^hr-*, and many others. As we have seen, consonants, individually and in clusters, bear most of the semantic weight of language sounds, having both strong indexical/[2] and symbolic/[3] properties, so it should not surprise that a language for which [+3] is the primary conditioning constraint would be “consonant heavy”. This phonotactic characteristic of Sanskrit (and other “consonant-heavy” languages) may be regarded as a syntagmatic manifestation of [+3] at the level of phonology.

The other distinctive typological trait of Sanskrit consonants is the consistent three-way set of featural oppositions among stops and affricates at every point of articulation from dental to velar. Sanskrit may add to any consonant either [+voice], [+aspiration], or both, resulting in four phonemic possibilities at each point of articulation, as shown on Table 16.

Table 16. Sanskrit stops and affricates

		[+asp]	[+voice]	[+voice, + asp]
Labial	p	p ^h	b	b ^h
Dental	t	t ^h	d	d ^h
Retroflex	ʈ	ʈ ^h	ɖ	ɖ ^h
Palatal	c	c ^h	j	j ^h
Velar	k	k ^h	g	g ^h

This state of affairs has the effect of modifying the fortis-lenis opposition (rather than altogether nullifying it, as Tamil has done). Instead of a binary opposition/[2] implied by the proliferation of fortis-lenis pairings (e.g., voiced/unvoiced stop phoneme pairs), Sanskrit presents a ternary array of complementary features, whereof fortis and lenis are represented as a graded series. Thus [+aspiration] would appear to be the most energetic (fortissimus), the voiceless, unaspirated version second most (fortior) and [+voiced] lenis. The status of [+asp, +voice] relative to this series is unclear; it would seemingly be intermediate between [+aspiration] and [+voice].

From a semiotic standpoint, though, there are three features (one the “zero” or unmarked feature, and the other two the additive features [+voice] and [+aspiration]). It would be fatuous to claim that any of the Categories per se are embodied in each of these features, since a) the nature of such features is possibly too vague to admit

of any such determination and b) Vedic Sanskrit being a “dead” language, a critical source of data, allophony, which we have made use of in our examination of other languages, is not available. It is sufficient to note that, even at this elemental level, Sanskrit establishes a triadic featural configuration, which triadic contours are suggestive of [3], instead of dyadic oppositions that typify the fortis-lenis featural axis, and that this triadic configuration is likely a manifestation of the conditioning feature [+3] in the paradigmatic realm of Sanskrit phonology.

We summarize the semiotic structures in Vedic Sanskrit conditioned by the constraint [+3] as follows:

[+3]:

Morphosyntax:

- Elaborate systems of subject-verb and noun-adjective agreement morphology
- Pervasive configuration of noun and verb paradigms according to trichotomies

Lexicon:

- Triadic canonical structure of many nouns and finite verbs (root-thematic element-affix)
- Ablaut trichotomy base form-guṇa-vṛddhi

Phonology:

- Very large variety of consonants and allowable consonant clusters
- Triadic featural configuration of obstruents

3. Conclusions

We have examined cases where [1], [12], [13], and [3] are manifest as conditioning constraints in languages, and have shown evidence that such Categorical constraints are detectable across various levels of language, imposing semiotic and structural consistency. This is, after all, the nature of all cognition qua Signs, to “reduce the manifold of sensuous impressions to unity”—that unity being the unity of a Sign. At the same time, we have deliberately chosen languages that happen to be familiar to this author, and that, by virtue of historical or geographical accident, or by cultural contrivance, appear to have been comparatively undiluted by foreign influences. Millenia ago, when geographical isolation was the norm, we must suppose that such comparative semiotic pristinity was also. But it is also the nature of Symbols to

evolve, which evolution frequently takes the form of engrafting new signs upon old. In the case of Sanskrit, for example, we note a number of changes asserting themselves as the Vedic language evolved into the epic and then the classical form. The idiom of the Vedic rishis possessed far greater inflectional complexity than that of Kālidāsa (not to mention descendant Prakrits like Pāli)—but the latter had at his disposal new, non-inflectional forms of morphosemiotic diversity that would have bewildered Angiras and his contemporaries. Learners of the classical language (normally the first version of Sanskrit learnt by aspiring Orientalists) are baffled by the extraordinary proliferation of synonyms in Classical Sanskrit, large numbers of which are, in effect, compound neologisms. Where the stylists of the Vedas seem to have held concision in high esteem, the opposite was true for the bards of the Mauryan and Guptan eras. A quick perusal of Monier-Williams' magisterial dictionary will disclose dozens of synonyms for words like "elephant", "lotus", "mountain", "bird", and so forth, not to mention hundreds of epithets for favorite deities like Shiva (most of whom were not featured members of the Vedic pantheon, but later innovations). Most synonyms that have the form of compounds appear to have arisen during the epic and classical eras, when, for some reason, the art of deploying and even inventing new synonyms was aesthetically prized. For example, the word for 'bird' in the Vedic language, *vi-*, is of obvious Indo-European origin (cf. Latin *avis*), but from the epic stage of Sanskrit onward, that word gave ground to a series of kenning-like neologisms (usually, though not always, compounds), such as *paksi-* (literally, 'winged'), *viha-ga-* ('sky-goer'), *nabhaścara-* ('sky-goer'), *nabhasamgama-* ('sky-goer'), *nāḍicarana-* ('stalk-legged'), *gaganacara-* ('moving in the air'), *dyuga-* ('sky-goer'), etc.

Sanskrit being in effect a refined, literary idiom, it is difficult to ascertain the degree to which the proliferation of synonyms and elaborate compounds were features of more demotic varieties of the language which have not come down to us. But it does appear to be at least plausible that, in the literary idiom at least, some of the characteristics of [13], the conditioning constraint of Tamil and the Dravidian semiotic universe, were beginning to penetrate Sanskrit. Certainly Dravidian borrowings become more and more commonplace in Sanskrit in later stages of the language, and there is no disagreement that the retroflex series of stop consonants in Sanskrit and its descendant Indo-Aryan languages were imported from Dravidian. But imported lexemes and phonemes are not, in the semiotic framework, mere meaningless objects; they are meaningful signs, representing something. And what do they represent? That

semiotic materials—i.e., meaningful structures—are being transferred from the source language to the recipient language. Given the wholesale importations of Dravidian lexemes into classical Sanskrit and subsequent Middle Indic, is it implausible to suppose that the spectacular profusion of neologisms—which, as we have seen, is typological feature of Tamil—also betokens a growing intermingling with Indo-Aryan of Tamil and Dravidian semiotic and aesthetic preferences?

In effect, relatively pristine, semiotically consistent languages like Mandarin, Tamil, and Vedic Sanskrit are the exception; the norm is the “booming, buzzing confusion” of modern languages, with their centuries, even millennia, of cross pollination and semiotic admixture. From such, it may be much more difficult to tease out semiotic regularities, without the uniformity of data afforded by such relatively undiluted semiotic systems as may still be accessible to the researcher. But having established the nature by which the conditioning constraints [+1], [+12], [+13], [+3], [-2], and [-3] may be manifest, both in the syntagmatic and paradigmatic domains in morphosyntax, in the lexicon, and in phonology (and, we hasten to add, the foregoing must be regarded as only a very modest beginning of such characterizations!), we are enabled to recognize them in their expected occurrence in other language, including languages embodying much more complex semiology begotten of linguistic and cultural intermingling. But such semiotic portraiture will have to await further research.

In the meantime, the methodology developed in the foregoing would appear to support the notions that semiotic structuring in human language 1) presents an identifiable consistency, as emblemized by the Categorical conditioning constraints we have identified (and probably others as well, such as [+23]), across different linguistic domains, 2) operates in both the syntagmatic and paradigmatic realms, and 3) is prior to all other structural schema (or, otherwise put, other structures arise in the first place from semiotic structures). Semiotic structures appear to be the true linguistic ultimates, the long-positied “deep structures”. We would suggest that their explanatory power is potentially broader than generative approaches, because the latter tend to assume different linguistic domains as modular, and typically assign ontological supremacy to syntax, whereas semiotic structures provide conceptual unity (the aim of any cognition, it cannot be too often repeated!) across the entire linguistic landscape.

The chief drawback of semiotic structures is, that in order for the investigator to compose a meaningful semiotic portrait of language, significant familiarity with

the language, including its cultural milieu and speaker community, is necessary, since many of the typologically significant traits drawn upon in studies like this (the proliferation of synonyms in Tamil or the near-universality of dilexic words in Chinese, for example) are not the sort of details likely to be evident in a typical descriptive grammar. Put more colloquially, this kind of work generally requires the researcher to become familiar enough with the language to truly “get a feel” for it, and not merely skim over data elicited by others in descriptive grammars. In such pursuits, the field anthropologist and the sociologist may enjoy a distinct advantage over the desk-bound theorist, because in the field of Peircean semiotics, the token of the realm is experience.

In contrast to theoretic approaches grounded in Cartesian dualism, the semiotic method embraces exceptions; traditional grammars—Cartesian contrivances all—may indeed leak, but semiotic structures accommodate and even expect variety. This is because all Signs include an element of Firstness, and it is Firstness, as we have seen, that begets variety, freshness, and diversity. The operability of “law” in the linguistic semiotic realm is evident in tendencies, preferences, and probabilities, not in exceptionless formulations. The proper posture with language, as Greenberg long ago recognized with implicational “universals”, is the semiotic validity of phenomena that are true “with overwhelmingly more than chance frequency”.

It may be objected that the Peircean Categories encompass so broad a range of characteristics as to be of little explanatory utility. We respond that, as vehicles for the reductionist approach so fashionable in modern science, they are not merely inadequate, but altogether incompatible. But our aim is to provide conceptual unity for the entirety of language, and for this, Cartesian reductionism, by requiring the often arbitrary neglect of the inconvenient messiness of human language, has had decidedly mixed explanatory results. The Peircean semiotic, with its all-encompassing posture, unifies formalism with aesthetic, discreteness with continuity, in a theoretical whole that contemplates language as a manifestation of semiotic universals that perfuse the natural world.

Notes

- 1 We shall hereafter use the term “Categorial” (as opposed to “categorical”) as the attributional term associated with the Categories. Moreover, note that words used in the (Peircean) semiotic sense will begin with a capital letter, to distinguish them from non-

semiotic senses; thus Category, Object, Sign, Event, Interpretant, etc., are understood to connote semiotic Category, semiotic Object, and so forth.

- 2 “[T]here can be no reality which has not the life of a symbol”, in “New Elements (*Kaina Stoicheia*)” (Peirce, 1998, p. 324).
- 3 Not all semiosis associated with cognition is explicit in the sense discussed here. Cognition consisting of thoughts and feelings, in addition to the explicit elements associated with human language, we define language here to consist of the subset of Thought-signs—words, paradigms, morphosyntax, etc.—selected for explicit representation by a particular community of Minds with a shared system for selective Representation.
- 4 Many Language-Signs and Culture-Signs are in fact hybrid, i.e., consist of more than one prominent Category-Sign. In this paper, however, we will confine our discussion to languages that appear to be conditioned predominantly by single Category-Signs.
- 5 As one example, consider a lexical peculiarity that has arisen as an inevitable outgrowth of the Chinese fascination with dichotomies: the extraordinary number of conjunctions in Chinese, especially conjunctions denoting the term ‘and’. Chinese has at least 11 in common usage, to wit: *hé, yǔ, ér, bìng, jí, gēn, érqiě, yìjǐ, bìngqiě, qiě*, and *jíqí*. Such a proliferation of conjunctions is a paradigmatic icon of the centrality in the Chinese semiotic universe of conjoining two terms, the essence of the dichotomy required by the constraint [+12].
- 6 Regarded as Signs-in-Themselves, all elements of the Subject are Indexes/[2], as Peirce pointed out (“The Categories Defended”, Peirce, 1998, pp. 172-173). We are considering here the way that each of these elements is represented as relating to the Predicate, not its manner of designating its Object.
- 7 Yes/no questions in Chinese may also be expressed by the addition of an interrogative like *ma* or *a* to a declarative sentence (such that, e.g., *tā yǒu-měi-yǒu* may alternatively be realized as *tā yǒu ma?*).
- 8 However, if the tones were removed but both consonants and vowels kept intact, such a passage would likely be as comprehensible as an English passage bereft of vowels, i.e., mostly comprehensible.
- 9 Tonality is a characteristic of many languages, and, while it is difficult to generalize about tonality across language groups, it is often the case that tones operate at the lexemic level, akin to their function in Chinese.
- 10 Du Ponceau wrote that polysynthetic languages express “the greatest number of ideas ... comprised in the least number of words. This is done principally in two ways. 1. By

a mode of compounding locutions which is not confined to joining two words together, as in the Greek, or varying the inflection or termination of a radical word as in the most European languages, but by interweaving together the most significant sounds or syllables of each simple word, so as to form a compound that will awaken in the mind at once all the ideas singly expressed by the words from which they are taken. 2. By an analogous combination of various parts of speech, particularly by means of the verb, so that its various forms and inflections will express not only the principal action, but the greatest possible number of the moral ideas and physical objects connected with it, and will combine itself to the greatest extent with those conceptions which are the subject of other parts of speech, and in other languages require to be expressed by separate and distinct words.... Their most remarkable external appearance is that of long polysyllabic words, which being compounded in the manner I have stated, express much at once” (Du Ponceau, 1819, pp. xxx-xxxii).

- 11 These three “persons” are transported by grammarians into other language groups, but in morphosemiotic terms, they are seldom valid. Consider, for example, that Tamil exhibits no less than 7 distinct morphological persons in the singular (first person, second person familiar, second person formal, third person masculine, third person feminine, third person formal, third person irrational) and four or five in the plural (first person inclusive, first person exclusive, second person, third person rational, third person irrational), and thus may be said to mark a first, second, and third person only by the Procrustean standards of European grammarians!
- 12 By the time of Classical Sanskrit, the perfect had evolved into a “historical past” tense, used only for events in the remote past that were not part of the writer’s personal experience.
- 13 Even athematic roots, which—not unlike Tamil “direct” noun forms—have no case affix in the nominative singular, nevertheless often appear in forms different from the root. For example, the root *rc-*, ‘verse’, is realized in the nominative singular as *rk* (*rg-* before case affixes beginning with a consonant, as, e.g., *rg-bhyah*, and *rc-* before case affixes beginning with a vowel (*rc-ah*).
- 14 Interestingly, the three Germanic ablaut grades in English, still manifest in English strong verbs like *swim* and *sing*, appear also to conform to the three Categories. Zero grade forms like *swim* and *sing* denote the present tense, which, as we have seen, is one of the cardinal manifestations of Firstness. The so-called “second principle part”, simple past stems like *swam* and *sang* denote the past tense—which is a Second. Participial forms

like *swum* and *sung* denote habitual action (as in *I have [often] sung*), and also serve to establish a relationship between past/[2] and present/[1] (as in *I have swum in this lake [before]*). Thus the ablaut marking the participial forms of English strong verbs may be said to signalize Thirdness, and the three English ablaut grades may be said to designate Firstness, Secondness, and Thirdness just as do the three Sanskrit grades. However, in English the picture is considerably less clear, inasmuch as some strong verbs only display two principle parts, and some roots display a fourth ablaut grade associated with nouns (as *sing* does, with *song*).

References

- Bonta, S. (2015). A Peircean typology of cultural prime symbols: Culture as category. *Semiotica*, (207), 251-277.
- Bonta, S. (2018). The lens of Firstness. *Semiotica*, (221), 143-174.
- Bonta, S. (2020). The Peircean sign as a model for interpreting semiotic structures in Mandarin Chinese. *Language and Semiotic Studies*, 6(1), 1-30.
- Deely, J. (1990). *Basics of semiotics*. Bloomington: Indiana University Press.
- Du Ponceau, P. S. (1819). Report of the corresponding secretary to the committee, of his progress in the investigation committed to him of the general character and forms of the languages of the American Indians. Read, 12th January, 1819. *Transactions of the Historical & Literary Committee of the American Philosophical Society, held at Philadelphia, for promoting useful knowledge 1*. pp. xvii–xlvi.
- Duanmu, S. (2013). How many Chinese words have elastic length? In F. Shi & G. Peng (Eds.), *Eastward flows the great river: Festschrift in honor of Prof. S.-Y. Wang's 80th Birthday* (pp. 1-14). Hong Kong: City University of Hong Kong Press.
- Jakobson, R., & Waugh, L. (1979). *The sound shape of language*. Bloomington: Indiana University Press.
- Kiparsky, P. (1998). Aspect and event structure in Vedic. In R. Singh (Ed.), *Yearbook of South Asian languages and linguistics*. Thousand Oaks: Sage Publications. Retrieved from <https://web.stanford.edu/~kiparsky/Papers/yearbook.pdf>
- Mithun, M. (1983). The genius of polysynthesis. In J. S. Thayer (Ed.), *North American Indians: Humanistic perspectives (University of Oklahoma Papers in Anthropology, Vol. 24)* (pp. 221-242). Norman: University of Oklahoma Press.
- Peirce, C. S. (1992). *The essential Peirce: Selected philosophical writings* (Vol. 1) (Peirce Edition Project, Eds.). Bloomington: Indiana University Press.

- Peirce, C. S. (1998). *The essential Peirce: Selected philosophical writings* (Vol. 2) (Peirce Edition Project, Eds.). Bloomington: Indiana University Press.
- Peirce, C. S. (1931-1958). *Collected papers of Charles Sanders Peirce* (Vols. 1-8) (C. Hartshorne, P. Weiss, & A. W. Burks, Eds.). Cambridge, MA: Harvard University Press. (cited as CP)
- Ramamurti, G. V. (1931). *A manual of the Sora (or Savara) language*. Madras: Government of Madras Press.
- Robertson, J. S. (1994). Linguistic implications of the Peircean sign and an analysis of the English inflectional morphemes -ing, -ed, and -s. *Semiotica*, 102(3/4), 179-223.

About the author

Steven Bonta (steve.c.bonta@gmail.com) received his PhD in linguistics from Cornell University (2004), and subsequently taught languages and linguistics at Penn State University from 2005-2016. He has published papers in semiotics and South Asian linguistics, and his areas of research include South Asian language contact (Sinhala and Tamil), Peircean semiotics, and writing systems of modern and ancient South Asia.