VLADIMIR PERICLIEV<br>Mathematical Linguistics Dpt., Institute of Mathematics \& Informatics, Bulgarian Academy of Sciences<br>peri@math.bas.bg

## ON REDUNDANCY IN DESCRIBING LINGUISTIC SYSTEMS


#### Abstract

The notion of system of linguistic elements figures prominently in most postSaussurian linguistics up to the present. A "system" is the network of the contrastive (or, distinctive) features each element in the system bears to the remaining elements. The meaning (valeur) of each element in the system is the set of features that are necessary and jointly sufficient to distinguish this element from all others. The paper addresses the problems of "redundancy", i.e. the occurrence of features that are not strictly necessary in describing an element in a system. Redundancy is shown to smuggle into the description of linguistic systems, this infelicitous practice illustrated with some examples from the literature (e.g. the classical phonemic analysis of Russian by Cherry, Halle, and Jakobson, 1953). The logic and psychology of the occurrence of redundancy are briefly sketched and it is shown that, in addition to some other problems, redundancy leads to a huge and unresolvable ambiguity of descriptions of linguistic systems (the Buridan's ass problem).


Keywords: redundancy; phonological and semantic systems; redundancy and ambiguity

## 1. Introduction

The notion of system of elements, in contrast to a mere agglomeratate of elements, figures prominently in most post-Saussurian linguistics up to the present. A "system" is the network of the contrasts each element in the system bears to the remaining elements, and as such the notion reflects the basic "logic of differentiation", a logic (knowledge) that every speaker of a language must be in firm possession of in order to be able to distinguish what is meaningful ( $=$ functional) from what is not in this language. And it is one of the few consensus points in contemporary linguistics that a major goal of linguistics is to uncover the various systems (phonological, grammatical or semantic) of individual languages.

The idea of system stems from Ferdinand de Saussure (Saussure, 1916/1996). Language (langue), Saussure taught, is a system in the sense that the meaning or value (valeur) of all linguistic entities can only be determined by their contrasts, distinctions from all other entities in the same system. "In the language itself, there are only differences", wrote Saussure (Saussure, 1916/1996, p. 118; italics in original), "A linguistic system is a series of phonetic differences matched with a series of conceptual differences" (Saussure, 1916/1996, p. 118). A central task of linguistics, according to Saussure, is to reveal the structure of linguistic systems by applying the structural method of contrasts and oppositions.

By definition, conforming to what we referred to above as the "logic of differentiation", each element within a system must be described ( $=$ profiled) in terms of all features needed for its discrimination from all other elements (which is the "sufficient condition for its discrimination"), and only them (which is the "necessary condition for its discrimination"). Thus, violating the sufficient condition for demarcation of an element will result in a failure to put apart this element from some other element within the system, while violating the necessary condition for its demarcation will introduce superfluousness, or redundancy, in the profiling of this element (and therefrom in our description of the system as a whole).

The notion of redundancy is generally connected to predictability. In language itself, redundancy represents an essential constitutive principle of communication, and is thus present in all semiotic systems. In language description, in contrast, redundancy has a negative connotation as it is usually associated with superfluity and overabundance in formal description, which should generally be avoided. This means that both grammar rules and lexicon rules should preferably be redundancyfree. The emergence of redundancy in the formulation of language universals was recently discussed in cases in which a universal is misstated, as it is actually a logical consequence of a stronger universal and is fully predictable from the latter (cf. Pericliev, 2012a).

In this paper, we explore the description of linguistic systems with regard to the emergence of redundancy in them. We show that redundancy does occur in actual descriptions of phonological and kinship systems and this creates problems. A particularly acute problem is the formidable ambiguity of redundant models, which cannot be readily resolved and is thus reminiscent of the problem of Buridan's ass, which, faced with two equidistant and desirable bales of hay, starves to death because there are no grounds for preferring to go to one bale rather than the other. There is still another harmful consequence from admission of redundancy, amounting to the creation of the mistaken view that redundant features are actually functional, or meaningful, for an element, which will impose illegitimate constraints on its permissible variability within the system.

The paper is organized as follows. Sections 2 and 3 are illustrations of redundant descriptions of linguistic systems in phonology. In Section 2, we show that the phonemic system of Russian, proposed by Cherry, Halle, and Jakobson (1953) is redundant, contrary to the proclaimed goal of these authors. We provide the correct (unique) nonredundant analysis of the system in terms of the same features Cherry, Halle, and Jakobson use. Section 3 discusses an example from a standard book on phonology (Spencer, 1996), showing that redundancy may creep into a
description even of a small dataset comprising only five vowels. Section 4 treats some problems arising from redundancy, focussing on the ambiguity of the resultant redundant systems. The huge number of possible alternative redundant models are shown and an attempt in componential analysis of kinship terminology to legitimize redundancy is critically examined. Section 5 examines some aspects of the logic and psychology of redundancy, or the reasons why redundancy is at all possible and why it can easily smuggle into linguistic description. The conclusion summarizes our contributions and sketches some implications of our results.

## 2. Example 1: The phonological system of Russian

In a seminal paper in Language, Cherry et. al. (1953) propose a phonemic system for Russian. Their express goal is to provide definitions of Russian phonemes with necessary and sufficient conditions and therefore eliminate all redundancy in the system (one of the ideas - aside from the purely linguistic considerations - being to satisfy the requirements of information theory for performing minimum number of steps for identification of each speech sound, an idea about language sounds incidentally abandoned later by Halle (personal communication)).

The authors describe the 42 Russian phonemes in terms of the following 11 binary features:
(1) vocalic
(2) consonantal
(3) compact
(4) diffuse
(5) grave
(6) nasal
(7) continuant
(8) voiced
(9) sharp
(10) strident
(11) stressed.

Table 1 gives the resultant componential analysis in terms of the 11 binary $(+/-)$ features used (ignore for the moment the brackets, so that all components, non-bracketed and bracketed, count). The notation in the table follows the IPA system of transcription, except in the following: a comma after a letter indicates palatalization; the accent mark is placed immediately before the vowel letter; and a strident stop is rendered by the same letter as the corresponding constrictive with the addition of a circumflex.

Table 1: Cherry et. al.'s componential model of Russian phonemes and its redundancy (redundant components enclosed in brackets).

|  | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $k$ | - | + | + |  | + |  | - | - | - |  |  |
| $k$, | - | + | $+$ |  | $+$ |  | - | - | + |  |  |
| $g$ | - | $+$ | $+$ |  | + |  | - | $+$ | - |  |  |
| $g$, | (-) | $+$ | $+$ |  | $+$ |  | - | + | + |  |  |
| $x$ | - | + | $+$ |  | $+$ |  | $+$ |  |  |  |  |
| $c$ | - | + | $+$ |  | - |  | - |  |  |  |  |
| J | - | $+$ | $+$ |  | - |  | $+$ | - |  |  |  |
| 3 | - | + | $+$ |  | - |  | $+$ | + |  |  |  |
| $t$ | - | $+$ | - |  | - | - | - | - | - | - |  |
| $t$, | - | $+$ | - |  | - | - | - | - | $+$ | - |  |
| $d$ | - | $+$ | - |  | - | (-) | - | $+$ | - | (-) |  |
| $d$, | (-) | $+$ | - |  | - | (-) | - | $+$ | + | (-) |  |
| $s$ | - | + | - |  | - | - | $+$ | - | - |  |  |
| $s$, | - | $+$ | - |  | - | - | $+$ | - | + |  |  |
| $z$ | - | $+$ | - |  | - | (-) | $+$ | $+$ | - |  |  |
| $z$, | - | $+$ | - |  | - | (-) | $+$ | $+$ | $+$ |  |  |
| $\hat{s}$ | - | $+$ | - |  | - | - | - | (-) | (-) | + |  |
| $n$ | - | $+$ | - |  | - | $+$ |  |  | - |  |  |
| $n$, | - | $+$ | - |  | - | + |  |  | + |  |  |
| $p$ | - | + | - |  | $+$ | - | - | - | - |  |  |
| $p$, | - | + | - |  | $+$ | - | - | - | + |  |  |
| $b$ | - | $+$ | - |  | $+$ | (-) | - | $+$ | - |  |  |
| $b$, | (-) | + | - |  | $+$ | (-) | - | + | $+$ |  |  |
| $f$ | - | + | - |  | $+$ | (-) | $+$ | - | - |  |  |
| $f$, | - | $+$ | - |  | $+$ | (-) | $+$ | - | + |  |  |
| $v$ | - | $+$ | - |  | $+$ | (-) | $+$ | $+$ | - |  |  |
| $v$, | - | + | - |  | $+$ | (-) | $+$ | $+$ | $+$ |  |  |
| $m$ | - | + | - |  | $+$ | $+$ |  |  | - |  |  |
| $m$, | - | + | - |  | $+$ | + |  |  | + |  |  |
| 'u | $+$ | - | - | $+$ | $+$ |  |  |  |  |  | + |
| $u$ | $+$ | - | - | $+$ | $+$ |  |  |  |  |  | - |
| 'o | $+$ | - | - | - | $+$ |  |  |  |  |  |  |
| 'e | $+$ | - | - | - | - |  |  |  |  |  |  |
| 'i | $+$ | - | - | $+$ | - |  |  |  |  |  | + |
| $i$ | + | - | - | $+$ | - |  |  |  |  |  | - |


|  | $(1)$ | $(2)$ | $(3)$ | $(4)$ | $(5)$ | $(6)$ | $(7)$ | $(8)$ | $(9)$ | $(10)$ | $(11)$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $a$ | + | - | + |  |  |  |  |  |  |  |
| $a$ | + | - | + |  |  |  |  |  |  |  | + |
| $r$ | + | + |  |  |  |  | - |  | - |  |  |
| $r$, | + | + |  |  |  |  | - |  | + |  |  |
| $l$ | + | + |  |  |  |  | + |  | - |  |  |
| $l$, | + | + |  |  |  |  | + |  | + |  |  |
| $j$ | - | - |  |  |  |  |  |  |  |  |  |

Consider the description of the phoneme $/ \mathrm{z} /$. The phoneme $/ \mathrm{z} /$ is defined in the table as the bundle of features [-vocalic $\&+$ consonantal $\&$-compact $\&-$ grave $\&-n a s a l \&+$ continuant $\&+$ voiced $\&-$ sharp]. The feature [-nasal] however makes no contribution to the demarcation of $/ \mathrm{z} /$ from the other phonemes in the data set since the set it discriminates, viz. $/ \mathrm{n} \mathrm{n}, \mathrm{mm} / /$, is also discriminated by the feature [+voiced]. (The palatal counterpart phoneme $/ \mathrm{z}, /$ is redundant for the same reason.)

Let us also consider the definition of the phoneme $/ \mathrm{g}$,/, which comprises the features [-vocalic \& + consonantal \& + compact \& + grave \& - continuant \& + voiced $\&+$ sharp]. The feature [-vocalic], however, is superfluous. It discriminates the set of phonemes /u'u o e i'i a'arr, ll,/. The subset /u'u o e i 'i a 'a/ is differentiated by the component [+consonantal], /ll,/ is differentiated by [ - continuant $], / \mathrm{r} /$ by [ + sharp] and /r,/ by [+voiced], that is, all phonemes discriminated by [-vocalic] are actually also discriminated by other features already used in the definition of $/ \mathrm{g}, /$. It thus turns out that the feature [-vocalic] is redundant and should be omitted.

Finally, an example of a phoneme whose definition contains two redundant components. The phoneme $/ \mathrm{b}, /$ is assigned the features $[-$ vocalic $\&+$ consonantal $\&$ -compact \& + grave \& - nasal \& - continuant \& + voiced \& +sharp]. The feature [+consonantal] discriminates the phonemes /u 'u o e i 'i a 'a/, the feature [-continuant] distinguishes $/ \mathrm{l} \mathrm{l}, /$, the feature $[+$ voiced] $/ \mathrm{r}, /$ and $[+$ sharp] $/ \mathrm{r}, \mathrm{l} /$, which comprises the set of all phonemes that [-vocalic] can discriminate. This circumstance makes the use of the latter in the definition of $/ b, /$ redundant. Besides, [-nasal] is also unnecessary, as the set $/ \mathrm{n} \mathrm{n}, \mathrm{mm} /$ / it distinguishes is also distinguished by the feature [+voiced].

The full list of redundancies is exhibited in Table 1 by enclosing the superfluous features in brackets.

The observation one can make from Table 1 is that there is a lot of redundancy in these authors analysis: the 17 feature values enclosed in brackets are those components which are superfluous for the demarcation. 11 out of a total of 42 phonemes have an incorrect description. As a result of this, the average number of features in a definition of a phoneme turns out to comprise 6.14, rather than 6.5 , components as suggested by Cherry et. al. (1953). Also, it is worth noting that some feature bundles are incorrect in more than one way; thus, the palatal phoneme $/ \mathrm{d}, /$ is assigned three superfluous feature values, and $/ \mathrm{d} /, / \hat{\mathrm{s}} /$, and the palatal /b,/ two such values.

## 3. Example 2: The English vowel system

It is interesting that not only large data sets, but also miniature ones can suffer from redundancy. By way of a simple phonological illustration, consider the data set comprising only five elements, viz. the vowels /i e a o u/. In the introductory book on phonology Phonology: Theory and Description, Spencer (1996, p. 124) presents an "underspecified matrix", i.e. "a matrix with all redundancies extracted" of English vowels in terms of the features 'back', 'high' and 'low', cf. Table 2. (Note that the place of redundant features is left blank by Spencer.)

Table 2: Spencer's redundant model of the system of five vowels /i, e, a, o, u/ (redundant components enclosed in brackets).

| Vowels | back | high | low |
| :---: | :---: | :---: | :---: |
| $i$ | - | + |  |
| $e$ | - | - | $(-)$ |
| $a$ |  |  | + |
| $o$ | + | - | - |
| $u$ | + | + |  |

However, a closer look at the analysis of Spencer shows that the feature bundle of the vowel /e/ is redundant, since the bundle [-back \& - high \& - low] contains the superfluous component [ - low]. Thus, this definition of $/ \mathrm{e} /$ is intended to discriminate most economically this phoneme from the remaining phonemes $/ \mathrm{i}$ a o u/. The feature [-back] in its definition distinguishes /e/from /a ou/, all of which are [+back], and the feature [-high] distinguishes $/ \mathrm{e} /$ from $/ \mathrm{u} /$, the last sound to be put apart (which is [+high]). The feature [-low] is obviously superfluous for the demarcation, contrary to what Spencer assumes, demonstrating that redundancy in discovering linguistic systems can smuggle even into very simple systems of interrelations. This requires some explanations, which will be given shortly. Before this, we turn to the problems engendered by the admission of redundancy.

## 4. Problems of redundancy: multiple solutions

There are several problems arising from the occurrence of redundancy in the description of linguistic systems, which are sketched below.

Descriptive inadequacy. One harmful consequence from admission of redundancy amounts to the creation of the mistaken view that redundant features are actually functional, or meaningful, for an element. This will impose illegitimate constraints on its permissible variability within the system (e.g. the English phoneme /e/ can actually have allophones which are not [-low], which the system description in Table 2 forbids).

Predictability. A redundant analysis is fully predictable from a nonredundant analysis, but not vice versa. This being the case, the usual requirement for economy (Occam's razor) would make the redundancy superfluous and hence omittable.

Ambiguity. If the above two problems are widely known, the problem of the introduction of ambiguity and indeterminacy by the admission of redundancy have not been sufficiently treated in the literature, and we turn to this below. We shall
discuss a case from kinship systems (by no means an isolated one), in which an attempt is made to legitimize redundant systems by reference to external, nonstructural reasons, viz. "psychological validity". It will be shown how to compute the number of all alternative redundant systems, given a nonredundant one, and it will be demonstrated that the number of alternatives is so huge that a rational choice of a single alternative is practically impossible (Buridan's ass problem).

Our illustrative example is Noricks (1987), a paper published in the authoritative journal American Anthropologist. Noricks undertakes, among other problems, the task of choosing a psychologically valid componential analysis of Niutao kinship terms (Niutao is a Polynesian language of Tuvalu). He tries to evaluate their cognitive validity, the measure of which is the ability to predict similarity judgments by Niutao native speakers in kin term triads tests. Noricks proposes four different overall features ( $=$ dimensions) that can partition the same data set, and within each of the four sets of alternative dimensions, comes up with two componential models: a nonredundant and a redundant one. Thus, he gets eight componential models in all, four nonredundant and four redundant. He then tests all eight models (using triad tests, estimating similarity judgments of subjects on kinship terms), reporting that one of the redundant models is significantly and consistently better supported by his experiments than the others.

We need not go into a detailed discussion of his result, but may focus on just one of his componential analyses. It employs the following semantic dimensions:
(1) Generation distance between ego and alter, with values
1.1 - more than one generation distant
1.2 - one generation distant
1.3 - zero generation distant
(2) Affinity of alter, with values
2.1 - consanguineal
2.2 - affinal
(3) Presence of a coeval consanguineal pair of opposite sex in the genealogical chain between ego and alter, with values
3.1 - present
3.2 - absent
(4) Sex of the connecting relative between ego and alter, with values
4.1 - female
4.2 - male
(5) Generation seniority of alter, with values
5.1 - senior
5.2 - junior
(6) Sex of alter, with values
6.1 - male
6.2 - female.

Noricks then presents (cf. Table 3) a componential scheme encompassing two componential analyses, one nonredundant, in which the only components that count are those given without brackets in the table, and one redundant, in which all the
components, with and without parentheses, count. Note that e1, e2, ...etc. in angle brackets denote the empty cells in one row and are given for later reference.

Table 3: Norick's redundant model of Niutao kinship terms (redundant components enclosed in parentheses and empty cells in angle brackets).

|  | Generation <br> $(1)$ | Affinity <br> $(2)$ | Oppos.Sex <br> $(3)$ | Sex 1st <br> $(4)$ | Seniority <br> $(5)$ | Alter Sex <br> $(6)$ |
| ---: | :---: | :---: | :---: | :---: | :---: | :---: |
| makupuna | 1.1 | $(2.1)$ | $<\mathrm{e} 1>$ | $<\mathrm{e} 2>$ | 5.2 | $<\mathrm{e} 3>$ |
| tupuna | 1.1 | $(2.1)$ | $<\mathrm{e} 1>$ | $<\mathrm{e} 2>$ | 5.1 | $<\mathrm{e} 3>$ |
| tamana | 1.2 | 2.1 | 3.2 | $<\mathrm{e} 1>$ | 5.1 | 6.1 |
| maatua | 1.2 | 2.1 | 3.2 | $<\mathrm{e} 1>$ | 5.1 | 6.2 |
| tama | 1.2 | 2.1 | 3.2 | $<\mathrm{e} 1>$ | 5.2 | $<\mathrm{e} 2>$ |
| tamatuangaane | 1.2 | 2.1 | 3.1 | 4.2 | 5.2 | $<\mathrm{e} 1>$ |
| maatuatuangaane | 1.2 | 2.1 | 3.1 | 4.2 | 5.1 | $(6.2)$ |
| tuaatina | 1.2 | 2.1 | 3.1 | 4.1 | $<\mathrm{e} 1>$ | $<\mathrm{e} 2>$ |
| fungaono | 1.2 | 2.2 | $<\mathrm{e} 1>$ | $<\mathrm{e} 2>$ | $<\mathrm{e} 3>$ | $<\mathrm{e} 4>$ |
| maa | 1.3 | 2.2 | 3.1 | $<\mathrm{e} 1>$ | $<\mathrm{e} 2>$ | $<\mathrm{e} 3>$ |
| aavanga | 1.3 | 2.2 | 3.2 | $<\mathrm{e} 1>$ | $<\mathrm{e} 2>$ | $<\mathrm{e} 3>$ |
| taina | 1.3 | 2.1 | 3.2 | $<\mathrm{e} 1>$ | $<\mathrm{e} 2>$ | $<\mathrm{e} 3>$ |
| tuangaane | 1.3 | 2.1 | 3.1 | $<\mathrm{e} 1>$ | $<\mathrm{e} 2>$ | $<\mathrm{e} 3>$ |

We should now look at the number of potential redundant componential models corresponding to the componential scheme of Table 3. How many are these models? To find this, we should, first, find the number of all potential redundant definitions of each kin term and, secondly, compute their product, since each redundant definition of a kin term can freely combine with all the redundant definitions of all other kin terms.

Let us turn to the first task, the computation of potentially redundant definitions of individual kin terms. Consider Noricks's redundant definition of the kin term makupuna (first row), presented with its numerical components: [1.1 \& (2.1) \& 5.2]. Here, Noricks marks one component, the one in brackets, viz. (2.1), as redundant. However, there are also three additional redundant components, viz. those employing dimensions Nos. 3, 4 or 6 (corresponding respectively to the empty cells $<\mathrm{e} 1\rangle,<\mathrm{e} 2\rangle$ and $<\mathrm{e} 3>$ in the table). (For the sake of simplicity, we assume here and in the following that all dimensions are applicable to all kin terms). Thus, having four potential redundant components - one given in parenthesis and three in angle brackets in the table - we can use them individually or combine them in pairs, triples, etc. to obtain further redundant definitions of makupuna, containing respectively one, two, three, etc. redundant components.

Let us now consider the alternative redundant definitions of makupuna containing just one redundant component. In addition to Noricks's definition [1.1 \& (2.1) $\& 5.2$ ], we will get the alternatives $[1.1 \&<\mathrm{e} 1>\& 5.2],[1.1 \&<\mathrm{e} 2>\& 5.2]$ and [1.1 \& <e3> \& 5.2], i.e. four alternatives in all.

We can proceed with definitions of the kin term makupuna with two redundant components. The possible two-way redundant combinations from four redundant components are six in number; hence the resultant definitions of the kin terms are $[1.1 \&(2.1) \&<\mathrm{e} 1>\& 5.2],[1.1 \&(2.1) \&<\mathrm{e} 2>\& 5.2]$, [1.1 \& (2.1) \& $5.2 \&$ $<\mathrm{e} 3>]$, $[1.1 \&<\mathrm{e} 1>\&<\mathrm{e} 2>\& 5.2$ ], $[1.1 \&<\mathrm{e} 1>\& 5.2 \&<\mathrm{e} 3>]$ and $[1.1 \&<\mathrm{e} 2>$ $\& 5.2 \&<\mathrm{e} 3>]$.

We can also form triples from these four elements, in which case the resultant redundant definitions of makupuna become four in number: [1.1 \& (2.1) \& <e1> $\&<\mathrm{e} 2>\& 5.2],[1.1 \&(2.1) \&<\mathrm{e} 1>\& 5.2 \&<\mathrm{e} 3>]$, $1.1 \&(2.1) \&<\mathrm{e} 2>\& 5.2$ $\&<\mathrm{e} 3>$ ], [1.1 \& <e $1>\&<\mathrm{e} 2>\& 5.2 \&<\mathrm{e} 3>]$.

Finally, we can form just one quadruple out of our four elements, yielding the kin term's definition with four redundant components: $[1.1 \&(2.1) \&<\mathrm{e} 1>\&$ $<\mathrm{e} 2>\& 5.2 \&<\mathrm{e} 3>$ ].

To rephrase the matter in more general and precise terms, the number of redundant definitions of a kin term depends on the number of potentially redundant components ( $=$ empty cells) that can be used in its definition (in the case under discussion, this number is four, corresponding to dimensions Nos. 2, 3, 4 and 6 ). To obtain the number of redundant definitions with one, two, three, etc. redundant components, the following "Combinations Formula", also referred to as "r-combination" or " $n$ choose $r$ ", is used:

$$
C(n, r)=n!/ r!(n-r)!
$$

where $n$ is the number of all potentially redundant components and $r$ is an n-way $(1,2, \ldots n)$ combination of these components. Thus, for instance, in the case of definitions with two redundant components (i.e. $n=4, r=2$ ), we will get:

$$
C(4,2)=4 \times 3 \times 2 \times 1 / 2 \times 1 \times(2 \times 1)=6
$$

All redundant definitions of makupuna, containing a different number of redundant components, should then be summed up to get the total number of such definitions for the kin term. This number is: $4+6+4+1=15$.

We need to find the number of redundant definitions of each kin term from Table 3. These are as follows (computed by the formula above):

## Table 4:

|  | No. of redundant |  |
| ---: | :---: | :---: |
|  | components | definitions |
| makupuna | 4 | 15 |
| tupuna | 4 | 15 |
| tamana | 1 | 1 |
| maatua | 1 | 1 |
| tama | 2 | 3 |
| tamatuangaane | 1 | 1 |
| maatuatuangaane | 1 | 1 |
| tuaatina | 2 | 3 |


|  | No. of redundant |  |
| ---: | :---: | :---: |
|  | components | definitions |
| fungaono | 4 | 15 |
| maa | 3 | 7 |
| aavanga | 3 | 7 |
| taina | 3 | 7 |
| tuangaane | 3 | 7 |

We can now proceed with the second step of our computation of all redundant componential models of Niutao kin terms in Table 3 by multiplying the numbers of redundant definitions of all kin terms: $15 \times 15 \times 1 \times 1 \times 3 \times 1 \times 1 \times 3 \times 15 \times 7 \times 7 \times 7 \times 7=$ 72930 375. This turns out to be a huge number.

In his article, Noricks argues that the redundant model presented in Table 3 is "a better performer" than the other seven models of the same data set he considers. We need not go into the details of his argumentation, but we cannot fail to ask why just this specific redundant model out of the 72930375 that we have calculated from his matrix? We could go further, and ask: What about the other redundant models, conforming to the other three dimension sets, which will have about as many alternatives? Apparently, in the face of so formidable a number of alternative redundant systems, we cannot in a rational way single out just one alternative by tests for psychological validity, since all alternatives should be similarly tested. In effect, Noricks's (and similar) attempts to legitimize redundancy in the description of kinship systems are unsuccessful. The same of course would apply to redundant phonological systems (not an uncommon trend in contemporary phonology).

## 5. The logic and psychology of redundancy

Let us first turn to the question of the logic of redundancy: Why is redundancy at all possible in ascribing a model to a data set (= language system)? The logic reason is the availability of more than one potential contrast between two system elements (i.e. when these elements are more "dissimilar"), and instead of using just one, the model profiles an element using two or more contrasting features (= components). No such possibility exists for a pair of elements that differ in just one feature-value (i.e. when these elements are "maximally similar"); this featurevalue must be used anyway in the model to achieve contrast. The second question relates to the psychology of redundancy, or to the question of why redundancy may smuggle into our analyses (as was the case in all our phonological and kinship illustrations). We believe that the emergence of such undesirable situations can be attributed to the computational complexity of the discrimination task, conforming to the redundancy-free requirement. The basic point is this. If we view human model-discovery as sequential acts of choosing a contrast between any two elements in the system, in order to avoid redundancy, at every point of this sequence at which alternatives exist, we must check all previous alternative decisions, such that are related to the current choice, and make the current choice in accordance with them, eventually revising our previous decisions. By way of illustration, let us try to see how the redundant analysis in Table 2 might have smuggled in. Suppose we
want to find the distinctive feature bundle of the phoneme $/ \mathrm{e} /$. Then its contrasts with the remaining elements in the system, viz. /i a o u/ are as follows:

Table 5: The contrasts of /e/ with other vowels

$$
\begin{array}{lll}
\text { (Step 1): } & \mathrm{e} \sim \mathrm{i}[- \text { high }] & \\
(\text { Step 2): } & \mathrm{e} \sim \mathrm{a}[- \text { back }] & \text { OR [-low] } \\
\text { (Step 3): } & \mathrm{e} \sim \mathrm{o}[- \text { back }] & \text { OR [-round] } \\
\text { (Step 4): } & \mathrm{e} \sim \mathrm{u}[- \text { back }] & \text { OR [-high }] \text { OR [-round }]
\end{array}
$$

Let us view the discrimination process of /e/ from all other phonemes as a sequence of decisions that need to be made. At Step 1, we have no choice and must use the component $[-$ high $]$. At Step 2, however, we do have a choice, between [-back] and [-low]. Suppose we choose [-low]. Up to this moment, /e/ is profiled with two components [-high, -low]. At step 3, we must add some of the features [-back] or [-round], since there is no other alternative for distinguishing /e/from /o/, so, adding, say, [-back], we get the definition [-high \& - low \& -back]. At Step 4, we do not need to add a component, since /e/ discriminates from /u/ by components already in the definition of /e/. This results in a superfluous [-low]. Why? The reason is that at Step 2 we should have memorized that we had made a choice between [-low] and [-back], and at Step3, we should have backtracked to this choice, and seen that we should have chosen not [-low], but [-back] instead, which is anyway needed for contrasting /e/ with the other vowels. This failure to revise our previous decision is thus the reason for the redundant component [ - low] to creep in. We should note, however, that such a process, in which one needs to keep record of all previous alternative decisions and be ready to revise them (called "non-monotonic") may be difficult to perform even by computers, let alone human analysts. This computational complexity of the task is the basic reason for the inconsistencies observed and discussed in detail of further kinship models proposed by human agents (cf. Pericliev, 2013).

## 6. Conclusion

In principle, all language systems (whether phonological or kinship) are potential candidates for assignment of a redundant model. As we have seen from our phonological and kinship semantics illustrations, this eventuality is not uncommonly realized in actual structural analyses of linguistic systems. Proposing a redundant model automatically legitimizes all other redundant models and opens up the back door to a legion of additional, and equally legitimate, redundant models over and above the formidable ambiguity already present in nonredundant analyses (for the multiplicity of solutions of kinship systems, cf. e.g. Pericliev, 2012b, 2013). This circumstance introduces the need to make a choice among a set of equally admissible redundant alternatives (Buridan's ass problem), but our linguistic task of making a rational choice, unlike Buridan's ass's, would further be aggravated by the usually immense number of alternatives.

Our considerations in this paper may be viewed in the context of two broad trends in contemporary linguistics, represented by those sticking to redundancy-
free descriptions and those propounding redundant descriptions for psychological, cognitive reasons. Our discussion has implications for both camps, though in different respects. Both camps should be aware of some difficulties inherent in their approaches: the theorists of the former camp must look for ways to ensure redundant-free descriptions as this is, as we have shown, a difficult task (and stick to the principle not only in words, but also in deeds), while the theorists of the latter camp must look hard for better solutions to handle the Buridan's ass problem than they have achieved so far.

## References

Burling, R. (1964). Cognition and componential analysis: God's truth or hocus-pocus? American Anthropologist, 66(1), 20-28. http://doi.org/10.1525/aa.1964.66.1. 02a00020
Cherry, E. C., Halle, M., \& Jakobson, R. (1953). Toward the logical description of languages in their phonemic aspect. Language, 29(1), 34-47. http://doi.org/10.2307/ 410451
Noricks, J. S. (1987). Testing for cognitive validity: Componential analysis and the question of extensions. American Anthropologist, 89 (2), 424-438. http://doi.org/10. 1525/aa.1987.89.2.02a00110
Pericliev, V. (2012a). Formulating and misformulating language universals. Folia Linguistica, $46(1), 211-232$. http://doi.org/10.1515/flin. 2012.7
Pericliev, V. (2012b). Machine componential analysis of Bulgarian kinship terminology and more on the problem of multiple solutions. Cognitive Studies / Études cognitives, 12, 231-251.
Pericliev, V. (2013). Componential analysis of kinship terminology. Basingstoke: Palgrave Macmillan. Retrieved from http://www.palgraveconnect.com/doifinder/10.1057/ 9781137031181
Saussure, F. de (1996). Course in general linguistics. (C. Bally \& A. Sechehaye (with A. Riedlinger), Eds., R. Harris, Trans.). Chicago: Open Court. (First published 1916)

Spencer, A. (1996). Phonology: Theory and description. Oxford: Blackwell Publishers.

## Acknowledgment

This work was supported by a core funding for statutory activities form the Institute of Mathematics and Informatics, Bulgarian Academy of Sciences.

The author declares that he has no competing interests.
This is an Open Access article distributed under the terms of the Creative Commons Attribution 3.0 PL License (http://creativecommons.org/licenses/by/3.0/pl/), which permits redistribution, commercial and non-commercial, provided that the article is properly cited.
(C) The Author 2015

Publisher: Institute of Slavic Studies, PAS, University of Silesia \& The Slavic Foundation

