

Journal of Universal Language 15-1
March 2014, 137-168

Componential Analysis of Kin Terms — Some Problems and their Solutions

Vladimir Pericliev

Institute of Mathematics and Informatics, Bulgaria

Abstract

Componential analysis of kinship vocabulary has a long tradition, but a look at the practicing of the method reveals two basic problems. The first pertains to ensuring “consistent” kin term definitions (i.e., definitions with necessary and sufficient components), and the second to the discovery and handling of multiple solutions to kinship systems. The paper introduces a computer program implemented to handle these two problems, and in general, designed to study kinship vocabulary. The program guarantees consistent componential models and the discovery of all alternative models, as well as introduces simplicity constraints to reduce the usually huge number of alternatives to a unique or just a couple of models. We illustrate the problems with examples from the literature and apply the program to a complex kinship vocabulary (Serbo-Croatian) to show its operation.

Keywords: componential analysis, kinship terms, the KINSHIP program

Vladimir Pericliev

Institute of Mathematics and Informatics, bl. 8 1113 Sofia, Bulgaria

Phone: (+3592)9792877; Email: peri@math.bas.bg

Received January 24, 2014; Revised February 20, 2014; Accepted February 27, 2014.

1. Introduction

The idea of “system” is fundamental in linguistics, and componential analysis is the method for exposing the systems of linguistic entities. In componential analysis, the meaning of the entities forming a system is described as a conjunction of smaller components that are necessary and jointly sufficient to distinguish each entity in the system from all others. Componential analysis is used at all levels of linguistic analysis: phonology (distinctive feature analysis), grammar, and semantics.

In the domain of kinship semantics, the method was introduced by Lounsbury, Goodenough, and others in the sixties of last century and was followed by numerous attempts to reveal the semantic structure of kin terms in various “exotic” languages, hoping to understand the meaning and use of the terms, and more optimistically, to highlight the categorization and world view of native speakers. The topic flourished for several decades, but as it happens all too often in science, after this peak the method became somewhat less visible in published work. Nevertheless the approach was not abandoned altogether: in linguistics, it continued to be quite regularly used in semantic analyses of various theoretical persuasions, and in anthropology, “formal analysis” (as the componential method is usually referred to in anthropology) continued to be an indispensable part of kin term studies. The semantic structure arrived at by componential analysis, as is well known, is important for constructing dictionary definitions, for translation purposes, and for historical reconstruction.

The componential analysis of kinship vocabulary has a long tradition, but a look at the practicing the method reveals two major problems. The first pertains to ensuring “consistent” kin term definitions (i.e., definitions with necessary and sufficient components), and the second to the discovery and handling of multiple solutions to kinship systems. The paper introduces a

computer program implemented to handle these two problems, and in general, designed to study kinship vocabulary. The program guarantees consistent componential models and the discovery of all alternative models, as well as introduces simplicity constraints to reduce the usually huge number of alternatives to a unique or just a couple of models. The operation of the program is illustrated with the analysis of the complex kinship system of Serbo-Croatian, a language that has not been previously subjected to componential analysis. A detailed discussion of the problems of componential analysis, the computer program and a sizable number of illustrative examples can be found in Pericliev (2013).

The paper is organized as follows. Section 2 introduces the basic notions and Section 3 sketches the problems of consistency of componential models and multiplicity of solutions by considering a couple of examples from the literature. Section 4 describes briefly our computer program, the features (dimensions) it uses, as well as its simplicity constraints intended to reduce the usually immense number of possible solutions. Section 5 lists the Serbo-Croatian kin term vocabulary, Section 6 lists one “componential scheme” of Serbo-Croatian, showing the formidable indeterminacy of possible componential models, and Section 7 explains how our simplicity constraints are applied to reduce this ambiguity to just two simplest componential models of the language. Section 8 concludes the discussion.

2. Componential Analysis of Kin Terms: Basic Notions

Componential analysis is a well-known method in linguistics. The procedure, as applied to kinship semantics, is familiar from the works of Goodenough (1956, 1967), Lounsbury (1956, 1964), Wallace & Atkins (1960), Hammel (1965), Leech (1974), and

more recently, Geeraerts (2010), and Bernard (2011). Here, we shall only briefly review some basic notions and then proceed with some analytical problems of the method, our computational machinery, and its application to Serbo-Croatian kinship terms.

The “kin terms” of a language, such as Serbo-Croatian *majka*, *otac*, *strina*, *unuk*, and so on are linguistic labels for a range of “kin types” (= denotata), which specify the genealogical position of one’s kin with respect to oneself. In the following, we shall use the following standard abbreviations (Murdock 1949) of atomic genealogical relationships in terms of which the kin types are expressed: Fa = ‘father,’ Mo = ‘mother,’ Br = ‘brother,’ Si = ‘sister,’ So = ‘son,’ Da = ‘daughter,’ Hu = ‘husband,’ and Wi = ‘wife.’ These atomic relationships are juxtaposed to express more distant kin types (relatives), as, for example, MoBr ‘mother’s brother,’ MoSi ‘mother’s sister,’ MoSiHu ‘mother’s sister’s husband,’ and so on.

The meaning of kin terms is represented by all kin types, or relatives, covered by the term. For example, the meanings of the Serbo-Croatian terms *strina* is FaBrWi (father’s brother’s wife) and *unuk* is SoSo (son’s son) and DaSo (daughter’s son). The set of all kin terms in a language is the “kinship vocabulary” of the language.

The basic goal of componential analysis is to determine the relevant conditions for distinguishing the meaning of any of the kin terms within the kinship vocabulary from any other. Put differently, componential analysis should find, for any kin term, the common features for all its attendant kin types, such that these common features demarcate this term from all other kin terms in the kinship vocabulary.

There are two major approaches within this general trend, the “social-structural” approach, and the “psychological” approach. The “psychological” approach is more ambitious and aims to reveal the native speakers’ intuitions and conception of the world of relatives. In contrast, the social-structural approach is only

intended to concisely and intelligibly summarize the kinship domain, correctly predicting the usage of kin terms in a linguistic community. That is, given any pair of individuals in a society, alongside with some (minimum) items of information about their relationship, such as their sex, generation, and so on from such a componential model the analyst should be able to infer whether they are kinfolk and what terms they would use to refer to one another. In this paper, we adopt the social-structural approach, which is prevalent to-date and is furthermore a necessary prerequisite for conducting the more ambitious psychological justifications of componential models.

As in most other grammatical tasks, a general common *adequacy requirement* to componential analysis would be to discover, for any data set, all and only the componential paradigms that describe the structure of the domain. This means that proposed componential models should be “consistent.” i.e., kin terms must be defined by necessary and sufficient features, and besides all componential models for a given data set should be revealed.

Additionally, as in other grammatical tasks, proposed componential models should be the *simplest*, as simplicity is a highly evaluated virtue of linguistic analyses. The requirement for parsimony in our case embodies two basic criteria: (i) choose the smallest number of contrasting features (dimensions) sufficient to describe the kinship vocabulary; and (ii) choose the smallest number of components in the definitions of every kin term in the vocabulary. We can also use two additional filtering devices if some kin terms allow of more than one definition even after the application of the above two criteria.

3. Problems of Componential Analysis of Kin Terms

A critical evaluation of the literature on componential analysis of kinship terminology reveals two common basic problems. Analysts have either failed to provide “consistent” componential models (i.e., models defining kin terms by necessary and sufficient features) or have failed to list all componential models of specific data sets. The first problem has generally remained unnoticed, while the second was emphasized by Robbins Burling (1964) and attracted considerable attention in the literature, but without any feasible solutions proposed to solve it. We address these problems in detail elsewhere and here we will only limit our discussion to short illustrations.

We start with the two conditions of consistency and then proceed with multiple solutions.

3.1. Inconsistency

3.1.1. Violation of Sufficiency Condition

The violation of a sufficiency condition means that a kin term cannot be properly demarcated from some other term(s) with the components listed in its definition.

In a paper appearing in *American Anthropologist*, Wordick (1973: 1249) makes the claim that English can be handled by just three dimensions, viz. (i) type, with values lineal vs. collateral vs. affinal, (ii) generation removal, with values 1, 0, -1, and (iii) sex of referent, with values male vs. female. This assertion, however, cannot be taken very seriously. Thus, from a first glance it is seen that we cannot distinguish between *father*, *grandfather*, *great grandfather*, and so on insofar as they are all lineal males of ascending generation; the same applies to the female counterparts

of these words, *mother*, *grandmother*, and *great grandmother*, which are all lineal females of ascending generation.

Things however cannot be remedied by simply letting the feature “generation” take further numbers as values. There is a problem with the first feature, “type.” for its values lineal, collateral, and affinal – under their usual interpretations – are not mutually exclusive, as feature values should be. Thus, relatives can be *both* collateral *and* affinal (e.g., those designating *uncle*, *aunt*, and so on according to Wallace & Atkins (1960: 61); Goodenough (1965: 285); Nogle (1974: 64), and elsewhere); under Goodenough’s conception (1965: 285) relatives may be both lineal and affinal (e.g., *father-in-law* and so on). A reinterpretation that does make these values mutually exclusive is also not a way out. For example, defining lineal and collateral to apply just to consanguineals, but not to affinals, we run into an impossibility to demarcate, say, *uncle* and *father*, for *uncle* would be neither lineal (obviously), nor collateral (having affinal kin types like FaSiHu), nor affinal (having consanguineal kin types like FaBr). So failing to achieve a contrast on the dimension of type, the terms will not contrast on the other two dimensions generation and sex as well, for which they have the same values.

Thus, Wordick’s proposal turns out to be substantially flawed in failing to satisfy the sufficiency condition of componential definitions.

3.1.2. Violation of Necessity Condition

The violation of a necessity condition means that a kin term definition contains superfluous, redundant components that are not strictly necessary for its demarcation from some other term(s).

We may consider as an example a case from the book by Nogle (1974) devoted to American English kinship vocabulary. Consider his definition of the kin term *father*: [nuclear, consanguineal, ascending generation, lineal, male]. The component “lineal”

however is redundant in this description. Thus, the feature “lineality” serves to discriminate *father* from the set of kin terms *brother*, *sister*, *uncle*, *aunt*, *cousin*, *nephew*, and *niece*, all of which are collateral. At the same time, the term *father* is already assigned the component “nuclear.” which distinguishes it from *uncle*, *aunt*, *cousin*, *nephew*, and *niece* (all of which are nonnuclear) and the component “ascending generation” distinguishes *father* from *brother* and *sister* (which are of contemporary generation).

Thus, the components “nuclear” and “contemporary generation.” collectively, discriminate all the words the component “lineal” demarcates, making the use of the latter superfluous. In addition to this example, there are further 16 kin terms in Nogle’s analysis, which are redundant, but these are treated elsewhere.

3.2. Multiple Solutions

Componential analysts have noted the problem of the existence of multiple componential models that can be arrived at analyzing the same data set. This problem was first raised in a famous paper by Burling (1964), entitled “Cognition and componential analysis: God’s truth or hocus pocus?.” Burling’s basic point in this article was to show the large number of logically possible alternative componential models of any given set of kin terms. Thus, if there are three items in the kin term set (call the items a, b, and c), one has three apparent choices: use a component which separates a from b and c; one which separates b from a and c; or one which separates c from a and b. The possibility of using components which are relevant for only a part of the set doubles the number of possibilities. Whereas for a three term set the possibilities number 6, for any given four kin term set they number 124. In general, the number of logically possible alternative componential models steeply increases with the increase of the number of kin terms that have to be discriminated. In Burling’s opinion, there are

no means to sensibly reduce this huge number of alternatives and the methods advocated are not equal to their goal — if this goal is to uncover “psychological validity” — and the conduct of componential analysis is therefore “hocus-pocus” rather than an enterprise that reveals “God’s truth.”

In his response to Burling’s challenge, Dell Hymes (1964) was less skeptic on the question, suggesting that further constraints from the culture and so on will limit the possibilities, but his short comment was also largely theoretical and did not introduce any definiteness in the argument. Other theorists in the ethnosemantic field (e.g., Goodenough 1967) seemed to agree on the seriousness of the problem of multiple solutions, but not much was done to really try to resolve the difficulty. Thus, the practitioners of componential analysis continued to provide models with little or no care at all about alternative analyses. Others attempted to further test a couple or so of alternative componential models with psychological methods (triad tests and so on) in an attempt to choose “psychologically valid” models, while still others substituted componential analysis with other methods (extensionists and so on) in search of better methods for representing native speakers’ categorization.

The review of the literature actually shows that the problem was never really understood in its depth, and the range of alternative solutions was not fully appreciated. One consensus point was however reached, and it was that it was not theoretical and programmatic studies, but analyses of concrete and exhaustive data sets that could resolve the multiple-solutions problem, and ultimately evaluate the utility of the componential analysis method. It became clear that whether one pursues the more ambitious goal of “psychological validity” (as in ethnosemantics) or the more modest goal of revealing structural relationships (as in lexical semantics and the present study), the method should be empirically tested to see whether it can produce a manageable number of alternatives or not.

Concluding, we should emphasize that both the problem of inconsistency and that of the failure to list all multiple solutions actually stem from the computational complexity of the task of componential analysis. Thus, first, the application of the method involves the need to make a large number of pair-wise contrasts between kin terms. For instance, if the analyzed language has 50 kin terms, one needs to examine $(50 \times 49) / 2 = 1225$ pairs (computed by the formula “n choose 2”). Secondly, the task of finding the necessary and sufficient features for demarcation is a well-known complex task of computer science (technically, NP-complete, cf. e.g., Valdes-Perez, Pereira, & Pericliev 2000). These circumstances force the conclusion that the task requires automation. Importantly, the needed computational tool should be able, in a principled manner, to reduce to manageable limits the formidable indeterminacy of completely unconstrained componential analysis.

4. The KINSHIP Program

4.1. An Overview of the Program

We have designed a computer program, called KINSHIP, whose aim is to try to resolve both the problem of consistency of componential models and to generate all alternative models, eventually trying to resolve the multiplicity of solutions problem.

The KINSHIP system is given as input the set of kin terms for a language with all their attendant kin types. Given this information, the system can generate all consistent componential models for the set, applying natural simplicity criteria to constrain the choices. KINSHIP is actually an extension of a sophisticated general-purpose class discrimination program that has found various applications in linguistics and outside of linguistics. Below

is only a brief overview of KINSHIP.

The program is endowed with a set of features (or dimensions) and with subroutines that determine, for each kin type, the value the kin type has for the inspected feature. For instance, the values of a kin type for the feature “sex” can be determined by the system by its last symbol (= link), knowing further the sex of all atomic relationships. Thus, the program can find that the kin type FaBr is sex = male, since its last link, viz. Br, is male, while FaSi is sex = female, since Si is female. The feature “generation” of a kin type is determined as a sum of the generations of the links constituting this kin type, where the latter are +1 for the parental relationships Fa and Mo, -1 for the filial relationships So and Da, and 0 for all remaining relationships Br, Si, Hu, and Wi; thus the program can compute that the kin type FaBr is generation = 1, since $+1 + 0 = 1$. Similarly, all feature values (components) are computed, and this is done for all kin types in the data set.

The program then transfers the components of the kin types into components of kin terms by finding those components that are possessed by *all* the kin types covered by a kin term. For example, a Serbo-Croatian kin term like *unuk* ‘grandson’ will have the components [sex = male, generation = -2] since both its kin types, viz. SoSo (son’s son) and DaSo (daughter’s son), are both of male sex of two generations below ego.

The program then proceeds with computing the dimensions of contrast, or contrasting features, that demarcate each kin term from each other term in the dataset. Having found these pair-wise contrasts between kin terms, the definition of each term is produced, which contains the necessary and jointly sufficient contrasting features (or components) that discriminate this term from all others in the dataset. The latter process is the most costly one in the algorithm of the program.

4.2. The Features Used by the Program

There are nearly twenty features defined in KINSHIP in order to handle languages of various types (Eskimo, Sudanese, Hawaiian, and so on). Below we list only those features that are used for the purposes of this paper. Some of these are self-explanatory, whereas others are explained by simple examples.

- (1) Generation of relative, with feature values
 - generation = 2
 - generation = 1
 - generation = 0
 - generation = -1
 - generation = -2, and so on

The value of the feature “generation” can be any integer, including a range of integers (bounded by \geq ‘equal or greater than’ or \leq ‘smaller or equal to’) to handle cross-generational kin terms.

- (2) Sex of relative, with feature values
 - sex = m — male
 - sex = f — female
- (3) Genealogical distance, with the integer feature values
 - distance = 1
 - distance = 2
 - distance = 3, and so on

This feature, analogously to “generation.” can take as value any integer, including a range of integers (bounded by \geq ‘equal or greater than’ or \leq ‘smaller or equal to’). The kin type Fa has the value distance = 1, FaBr has the value distance = 2, and FaBrWi

distance = 3.

- (4) Affinity of relative, with two feature values
- affinity = cons – consanguineal, or absence of a marital tie
 - affinity = aff – affinal, marital tie

For example, the kin type MoSi has the value affinity = cons, while Wi and SiHu have the feature value affinity = aff.

- (5) Affinity of the 1st connecting relative (link), with feature values
- affinity 1st link = cons – consanguineal (first link is a blood relative)
 - affinity 1st link = aff – affinal (first link is a relative by marriage)

For example, the kin types FaSi, SiHu (where first link is underlined) both have the feature value affinity 1st link = cons, while WiMo, HuSi are both affinity 1st link = aff.

- (6) Sex of the 1st connecting relative (link), with feature values
- sex 1st link = m – male
 - sex 1st link = f – female

For example, the kin types Fa and FaSi (where first link is underlined) both have the feature value sex 1st link = m, while Mo and WiSi are both sex 1st link = f.

- (7) Generation of the last link, with the feature values
- generation last link = 1
 - generation last link = 0
 - generation last link = -1

The generations of the last links will be: Fa and Mo = 1; Si, Br, Hu, and Wi = 0; So and Da = -1; hence, e.g., FaMo (where last link is underlined) will have the feature value generation last link = 1 for this feature, while So or SoSo will have the feature value generation last link = -1.

- (8) Sex of the second connecting relative (link), with feature values
- sex 2nd link = m — male
 - sex 2nd link = f — female

For example, the kin types FaFaSi, MoBr (where second link is underlined) both have the feature value sex 2nd link = m, while FaMoSi, MoSi are both sex 2nd link = f.

- (9) Affinity of the last connecting relative (link), with feature values
- affinity last link = cons — consanguineal
 - affinity last link = aff — affinal

For example, the kin types FaFaBr, MoBr (where the last link is underlined) both have the feature value affinity last link = cons, while FaBrWi, SiHu are both affinity last link = aff.

- (10) Lineality, with the feature values
- lineality = lin — lineal
 - lineality = coll — collateral
- (11) Seniority within one generation, with feature values
- seniority = e — elder
 - seniority = y — younger

4.3. The Simplicity Constraints of the Program

KINSHIP uses four intuitive criteria to guarantee the uncovering of the simplest discrimination of the kinship terms. They refer to dimensions and components in kin term definitions.

(12) Constraints on dimension sets

a. *Minimize overall features (= dimensions).*

A set of kin terms may be demarcatable, using a number of overall feature sets of different cardinality; this criterion chooses those overall feature sets which have the smallest cardinality (i.e., are the shortest).

b. *Minimize “less universal” features in dimension sets.*

This criterion minimizes the use of “less universal” features in dimension sets, in the case when alternatives remain even after the application of the first criterion. The basic idea is that in the alternative, equally short dimension sets, some features are *obligatory* and must be used, while others are *optional* and may or may not be used. From the optional features in dimension sets, the criterion prefers the sets with “more universal” features, i.e., those that are known as far back as Kroeber (1909) to be universally used in the languages of the world (sex of relative, sex of speaker, generation, affinity, genealogical distance, lineality, seniority within one generation).

(13) Constraints on features (= components) in kin term definitions

c. *Minimize features (= components) in kin term definitions.*

Given some overall feature set, one kin term may be demarcatable — using only features from this set — by a number of definitions of different cardinality; this criterion

chooses those definitions, having the smallest cardinality (i.e., are the shortest).

d. *Minimize “minor” features in kin term definitions.*

This criterion minimizes the use of “minor.” i.e., infrequent, peripheral features in kin term definitions, in the case when alternatives remain even after the application of the previous three simplicity criteria. The basic idea is that in the alternative, equally short definitions, of one kin term, some features are *obligatory* and must be used, while others are *optional* and may or may not be used. From the optional features for one kin term, the criterion prefers the more frequent feature/s from a frequency hierarchy computed from the obligatory features used in the whole componential scheme (see below).

5. The Serbo-Croatian Kin Terms Data

Serbo-Croatian is an Indo-European language, of the south Slavic branch, and is the primary language of Serbia, Croatia, Bosnia and Herzegovina, and Montenegro. It is a pluricentric language having four mutually intelligible standard varieties. We have chosen Serbo-Croatian as an illustrative example for several reasons. First, it is a language which has not been previously analyzed in componential terms (most attempts in this direction address more “exotic” and less studied languages). Secondly, Serbo-Croatian is basically of a Sudanese type and has quite a complex system of kinship terminology, which is a challenge to semantic analysis. And, thirdly, Serbo-Croatian nicely illustrates the problem of multiplicity of solutions and how our program handles it.

Below we list the 51 Serbo-Croatian kin terms with their

attendant kin types. The list is compiled from various dictionaries and other sources and is pretty complete for our illustrative purposes, without pretending to be exhaustive. Completely excluded are the step-relatives and the foster-relatives. In compliance with common practice, we generally do not include terms with intersecting kin types or terms such that included / are included into others, as they do not cleanly discriminate (= do not form a “paradigm”).

Table 1. Serbo-Croatian Kin Terms with their Attendant Kin Types

| | |
|-------------------------|-----------------------------|
| <i>praded</i> | FaFaFa FaMoFa MoFaFa MoMoFa |
| <i>prababa</i> | FaFaMo FaMoMo MoFaMo MoMoMo |
| <i>djed</i> | FaFa MoFa |
| <i>baba</i> | FaMo MoMo |
| <i>otac</i> | Fa |
| <i>majka</i> | Mo |
| <i>stric</i> | FaBr |
| <i>strina</i> | FaBrWi |
| <i>ujak</i> | MoBr |
| <i>ujna</i> | MoBrWi |
| <i>tetka</i> | MoSi FaSi |
| <i>tetak</i> | MoSiHu FaSiHu |
| <i>brat</i> | Br |
| <i>sestra</i> | Si |
| <i>brat od strica</i> | FaBrSo |
| <i>brat od ujaka</i> | MoBrSo |
| <i>brat od tetke</i> | FaSiSo MoSiSo |
| <i>sestra od strica</i> | FaBrDa |
| <i>sestra od ujaka</i> | MoBrDa |
| <i>sestra od tetke</i> | FaSiDa MoSiDa |
| <i>bratić</i> | MoSiSo MoBrSo FaSiSo FaBrSo |
| <i>sestrična</i> | MoSiDa MoBrDa FaSiDa FaBrDa |
| <i>sin</i> | So |

| | |
|--------------------|-----------------------------|
| <i>ćerka</i> | Da |
| <i>nećak</i> | BrSo SiSo |
| <i>nećakinja</i> | BrDa SiDa |
| <i>bratanić</i> | BrSo |
| <i>brataničina</i> | BrDa |
| <i>sestrić</i> | SiSo |
| <i>sestričina</i> | SiDa |
| <i>unuk</i> | SoSo DaSo |
| <i>unuka</i> | DaDa SoDa |
| <i>praunuc</i> | SoSoSo SoDaSo DaSoSo DaDaSo |
| <i>praunuka</i> | DaDaDa SoSoDa SoDaDa DaSoDa |
| <i>maž</i> | Hu |
| <i>žena</i> | Wi |
| <i>svekar</i> | HuFa |
| <i>svekarva</i> | HuMo |
| <i>punac</i> | WiFa |
| <i>punica</i> | WiMo |
| <i>zet</i> | SiHu DaHu |
| <i>snaha</i> | BrWi SoWi |
| <i>šurjak</i> | WiBr |
| <i>šurjakinja</i> | WiBrWi |
| <i>šogorica</i> | WiSi |
| <i>dever</i> | HuBr |
| <i>badženjak</i> | WiSiHu |
| <i>jetrva</i> | HuBrWi |
| <i>zaova</i> | HuSi |
| <i>prija</i> | SoWiMo DaHuMo |
| <i>prijatelj</i> | SoWiFa DaHuFa |

6. The Alternative Componential Models of Serbo-Croatian

The KINSHIP system was run on the Serbo-Croatian dataset with the features described in Section 4, requiring that all alternative componential models be found. The system reported nine alternative overall feature (= dimension) sets that are necessary and jointly sufficient to demarcate all Serbo-Croatian kin terms. These overall feature (dimension) sets are listed below in terms of their abbreviations from Section 4, alongside with the number of features each of these sets comprises (rightmost column).

Table 2. All Alternative Feature (Dimension) Sets for Serbo-Croatian and their Sizes (Rightmost Column)

| No | Feature (dimension) sets | No |
|------|---|----|
| I | {sex 2nd link, sex 1st link, sex, generation last link, generation 1st link, distance, affinity 1st link} | 7 |
| II | {sex 2nd link, sex 1st link, sex, generation last link, generation, distance, affinity 1st link} | 7 |
| III | {sex 2nd link, sex 1st link, sex, generation, affinity, distance, affinity 1st link} | 7 |
| IV | {sex 2nd link, sex 1st link, sex, generation last link, distance, generation 1st link, affinity last link} | 7 |
| V | {sex 2nd link, sex 1st link, sex, generation 1st link, distance, generation, affinity last link} | 7 |
| VI | {sex 2nd link, sex 1st link, sex, generation last link, distance, generation, affinity last link} | 7 |
| VII | {sex 2nd link, sex 1st link, sex, affinity, distance, generation, affinity last link} | 7 |
| VIII | {sex 2nd link, sex 1st link, sex, generation last link, lineality, affinity, distance, affinity 1st link} | 8 |
| IX | {sex 2nd link, sex 1st link, sex, generation last link, lineality, distance, affinity 1st link, affinity last link} | 8 |

To each overall feature set, there correspond a huge number of componential models, resulting from the different definitions that can be given to some kin terms, using only features from this overall feature (= dimension) set. We call such a concise representation of multiple componential models a “componential scheme.” For economy of space, below we give just one componential scheme produced by our program, viz. that corresponding to the third dimension set, viz. {sex 2nd link, sex 1st link, sex, generation, affinity, distance, affinity 1st link}.

Table 3. Componential Scheme of Serbo-Croatian with the Third Dimension Set {sex 2nd link, sex 1st link, sex, generation, affinity, distance, affinity 1st link}

praded

1 [sex=m, generation=3]

prababa

1 [sex=f, generation=3]

djed

1 [sex 2nd link=m, generation=2]

2 [sex=m, generation=2]

baba

1 [sex 2nd link=f, generation=2]

2 [sex=f, generation=2]

otac

1 [sex 1st link=m, generation=1, distance=1]

2 [sex=m, generation=1, distance=1]

majka

1 [sex 1st link=f, generation=1, distance=1]

2 [sex=f, generation=1, distance=1]

stric

1 [sex 1st link=m, affinity 1st link=cons, sex 2nd link=m, generation=1, distance=2]

2 [sex 1st link=m, affinity 1st link=cons, sex=m, generation=1, distance=2]

- 3 [sex 1st link=m, sex 2nd link=m, generation=1, distance=2, affinity=cons]
- 4 [sex 1st link=m, sex=m, generation=1, distance=2, affinity=cons]

strina

- 1 [sex 1st link=m, sex 2nd link=m, sex=f, generation=1]
- 2 [sex 1st link=m, sex 2nd link=m, generation=1, distance=3]
- 3 [sex 1st link=m, sex=f, generation=1, distance=3]
- 4 [sex 1st link=m, affinity 1st link=cons, sex 2nd link=m, generation=1, affinity=aff]
- 5 [sex 1st link=m, affinity 1st link=cons, sex=f, generation=1, affinity=aff]

ujak

- 1 [sex 1st link=f, affinity 1st link=cons, sex 2nd link=m, sex=m, generation=1]
- 2 [sex 1st link=f, affinity 1st link=cons, sex 2nd link=m, generation=1, distance=2]
- 3 [sex 1st link=f, affinity 1st link=cons, sex=m, generation=1, distance=2]
- 4 [sex 1st link=f, sex 2nd link=m, generation=1, distance=2, affinity=cons]
- 5 [sex 1st link=f, sex=m, generation=1, affinity=cons]

ujna

- 1 [sex 1st link=f, sex 2nd link=m, generation=1, distance=3]
- 2 [sex 1st link=f, sex=f, generation=1, distance=3]
- 3 [sex 1st link=f, affinity 1st link=cons, sex 2nd link=m, generation=1, affinity=aff]
- 4 [sex 1st link=f, affinity 1st link=cons, sex=f, generation=1, affinity=aff]
- 5 [sex 1st link=f, sex 2nd link=m, sex=f, generation=1, affinity=aff]

tetka

- 1 [affinity 1st link=cons, sex 2nd link=f, generation=1, distance=2]
- 2 [affinity 1st link=cons, sex=f, generation=1, distance=2]
- 3 [sex 2nd link=f, generation=1, distance=2, affinity=cons]
- 4 [sex=f, generation=1, distance=2, affinity=cons]

tetak

- 1 [sex 2nd link=f, generation=1, distance=3]
- 2 [sex=m, generation=1, distance=3]
- 3 [affinity 1st link=cons, sex=m, generation=1, affinity=aff]

4 [affinity 1st link=cons, sex 2nd link=f, generation=1, affinity=aff]

5 [sex 2nd link=f, sex=m, generation=1, affinity=aff]

brat

1 [sex 1st link=m, affinity 1st link=cons, generation=0, distance=1]

2 [sex 1st link=m, generation=0, distance=1, affinity=cons]

3 [affinity 1st link=cons, sex=m, generation=0, distance=1]

4 [sex=m, generation=0, distance=1, affinity=cons]

sestra

1 [sex 1st link=f, affinity 1st link=cons, generation=0, distance=1]

2 [affinity 1st link=cons, sex=f, generation=0, distance=1]

3 [sex 1st link=f, generation=0, distance=1, affinity=cons]

4 [sex=f, generation=0, distance=1, affinity=cons]

brat od strica

1 [sex 1st link=m, sex 2nd link=m, sex=m, generation=0, distance=3, affinity=cons]

brat od ujaka

1 [sex 1st link=f, sex 2nd link=m, sex=m, generation=0, affinity=cons]

brat od tetke

1 [sex 2nd link=f, sex=m, generation=0, distance=3, affinity=cons]

sestra od strica

1 [sex 1st link=m, sex 2nd link=m, sex=f, generation=0, affinity=cons]

sestra od ujaka

1 [sex 1st link=f, sex 2nd link=m, sex=f, generation=0, distance=3, affinity=cons]

sestra od tetke

1 [sex 2nd link=f, sex=f, generation=0, distance=3, affinity=cons]

bratić

1 [sex=m, generation=0, distance=3, affinity=cons]

sestrična

1 [sex=f, generation=0, distance=3, affinity=cons]

sin

1 [sex 1st link=m, generation= -1, distance=1]

2 [sex=m, generation= -1, distance=1]

ćerka

1 [sex 1st link=f, generation= -1, distance=1]

2 [sex=f, generation= -1, distance=1]

nećak

1 [sex 2nd link=m, generation= -1, distance=2, affinity=cons]

2 [sex=m, generation= -1, distance=2, affinity=cons]

nećakinja

1 [sex 2nd link=f, generation= -1, distance=2, affinity=cons]

2 [sex=f, generation= -1, distance=2, affinity=cons]

bratanić

1 [sex 1st link=m, sex 2nd link=m, generation= -1, distance=2]

2 [sex 1st link=m, sex=m, generation= -1, distance=2]

brataničina

1 [sex 1st link=m, sex=f, generation= -1, affinity=cons]

2 [sex 1st link=m, sex 2nd link=f, generation= -1, distance=2, affinity=cons]

sestrić

1 [sex 1st link=f, sex 2nd link=m, generation= -1, distance=2, affinity=cons]

2 [sex 1st link=f, sex=m, generation= -1, affinity=cons]

sestričina

1 [sex 1st link=f, sex 2nd link=f, generation= -1, distance=2]

2 [sex 1st link=f, sex=f, generation= -1, distance=2]

unuk

1 [sex 2nd link=m, generation= -2, affinity=cons]

2 [sex=m, generation= -2, affinity=cons]

unuka

1 [sex 2nd link=f, generation= -2, affinity=cons]

2 [sex=f, generation= -2, affinity=cons]

praunuc

1 [sex=m, generation= -3, distance=3]

2 [sex=m, generation= -3, affinity=cons]

praunuka

1 [sex=f, generation= -3, distance=3]

2 [sex=f, generation= -3, affinity=cons]

maž

1 [sex 1st link=m, affinity 1st link=aff, distance=1]

2 [sex 1st link=m, distance=1, affinity=aff]

3 [affinity 1st link=aff, sex=m, distance=1]

4 [sex=m, distance=1, affinity=aff]

žena

1 [sex 1st link=f, affinity 1st link=aff, distance=1]

2 [affinity 1st link=aff, sex=f, distance=1]

3 [sex 1st link=f, distance=1, affinity=aff]

4 [sex=f, distance=1, affinity=aff]

svekar

1 [sex 1st link=m, affinity 1st link=aff, sex 2nd link=m, generation=1]

2 [sex 1st link=m, sex 2nd link=m, generation=1, distance=2, affinity=aff]

3 [sex 1st link=m, affinity 1st link=aff, sex=m, generation=1]

4 [sex 1st link=m, sex 2nd link=m, sex=m, generation=1, affinity=aff]

5 [sex 1st link=m, sex=m, generation=1, distance=2, affinity=aff]

svekarva

1 [sex 1st link=m, affinity 1st link=aff, sex 2nd link=f, generation=1]

2 [sex 1st link=m, affinity 1st link=aff, sex=f, generation=1]

3 [sex 1st link=m, sex 2nd link=f, sex=f, generation=1, affinity=aff]

4 [sex 1st link=m, sex 2nd link=f, generation=1, distance=2, affinity=aff]

5 [sex 1st link=m, sex=f, generation=1, distance=2, affinity=aff]

punac

1 [sex 1st link=f, affinity 1st link=aff, sex 2nd link=m, generation=1]

2 [sex 1st link=f, sex 2nd link=m, generation=1, distance=2, affinity=aff]

3 [sex 1st link=f, affinity 1st link=aff, sex=m, generation=1]

4 [sex 1st link=f, sex 2nd link=m, sex=m, generation=1, affinity=aff]

5 [sex 1st link=f, sex=m, generation=1, distance=2, affinity=aff]

punica

1 [sex 1st link=f, affinity 1st link=aff, sex 2nd link=f, generation=1]

2 [sex 1st link=f, affinity 1st link=aff, sex=f, generation=1]

3 [sex 1st link=f, sex 2nd link=f, sex=f, generation=1, affinity=aff]

4 [sex 1st link=f, sex 2nd link=f, generation=1, distance=2, affinity=aff]

5 [sex 1st link=f, sex=f, generation=1, distance=2, affinity=aff]

zet

- 1 [sex 1st link=f, affinity 1st link=cons, distance=2, affinity=aff]
- 2 [affinity 1st link=cons, sex 2nd link=m, distance=2, affinity=aff]
- 3 [affinity 1st link=cons, sex=m, distance=2, affinity=aff]

snaha

- 1 [sex 1st link=m, affinity 1st link=cons, distance=2, affinity=aff]
- 2 [affinity 1st link=cons, sex 2nd link=f, distance=2, affinity=aff]
- 3 [affinity 1st link=cons, sex=f, distance=2, affinity=aff]

šurjak

- 1 [sex 1st link=f, affinity 1st link=aff, sex 2nd link=m, sex=m, generation=0]
- 2 [sex 1st link=f, affinity 1st link=aff, sex 2nd link=m, generation=0, distance=2]
- 3 [sex 1st link=f, affinity 1st link=aff, sex=m, generation=0, distance=2]

šurjakinja

- 1 [sex 1st link=f, affinity 1st link=aff, sex 2nd link=m, distance=3]
- 2 [sex 1st link=f, affinity 1st link=aff, sex=f, distance=3]

šogorica

- 1 [sex 1st link=f, sex 2nd link=f, generation=0, distance=2]
- 2 [sex 1st link=f, sex=f, generation=0, distance=2]

dever

- 1 [sex 1st link=m, sex 2nd link=m, generation=0, distance=2]
- 2 [sex 1st link=m, sex=m, generation=0, distance=2]

badženjak

- 1 [sex 1st link=f, affinity 1st link=aff, sex 2nd link=f, sex=m]
- 2 [affinity 1st link=aff, sex 2nd link=f, distance=3]
- 3 [affinity 1st link=aff, sex=m, distance=3]

jetrva

- 1 [sex 1st link=m, affinity 1st link=aff, sex 2nd link=m, sex=f]
- 2 [sex 1st link=m, affinity 1st link=aff, distance=3]

zaova

- 1 [sex 1st link=m, affinity 1st link=aff, sex 2nd link=f, sex=f, generation=0]
- 2 [sex 1st link=m, affinity 1st link=aff, sex 2nd link=f, generation=0, distance=2]
- 3 [sex 1st link=m, affinity 1st link=aff, sex=f, generation=0,

distance=2]
prija
 1 [affinity 1st link=cons, sex=f, generation=0, distance=3, affinity=aff]
prijatelj
 1 [affinity 1st link=cons, sex=m, generation=0, distance=3,
 affinity=aff]

In a componential scheme, the different definitions of one kin term can freely combine with any other alternative definition of other terms to produce a componential model. As a consequence, the total number of alternative componential models, M^a , represented by a componential scheme are equal to the product of the number of definitions, N , each individual term has obtained, expressed by the formula $M^a = N1 \times N2 \times N3 \dots \times Nm$. For example, assuming our dataset to comprise only three terms, the first having two definitions, the second one, and the third four, we have $M^a = 2 \times 1 \times 4 = 8$ componential analyses in all. In our particular case in Table 3, there are eighteen terms having two definitions, four terms having three definitions, six terms having four definitions, and eight terms having five definitions, i.e., we get $2^{18} \times 3^4 \times 4^6 \times 5^8 = 33,973,862,400,000,000$. This is a huge number of componential models and so are the models corresponding to the other dimension sets, which we do not need to calculate here. The total number of alternative componential analyses is the sum total of all these and would be an astronomical number.

The only conclusion we can draw from this formidable indeterminacy is that the completely unconstrained componential analysis, in which the adequacy criterion is to provide all and only the “consistent” models corresponding to a kin term data set, leads to an intolerably large number of models. Without some constraints which limit the indeterminacy in some principled manner the method itself is in jeopardy, as Burling (1964) had warned (see the discussion in section 3.2). We therefore need to

apply our simplicity restrictions to the analysis of Serbo-Croatian, which we do in the next section.

7. Simplicity Constraints on Alternative Componential Models

Our simplicity constraints pertain to alternatives in dimensions and in kin term definitions and we discuss these in sequence.

7.1. Constraints on Dimensions

Our first constraint on dimensions chooses the dimension set(s) comprising the smallest number of features (= dimensions). Looking at Table 2, we observe that from a total number of nine dimension sets that are necessary and sufficient to demarcate all Serbo-Croatian kin terms, the first seven contain 7 features, and the last two have 8 features. So, in accord with our first criterion, we discard the last two dimension sets.

A lot of alternative dimension sets still remain (seven in number) and we recourse to our second criterion. It amounts to preferring dimension sets with features that have a more universal nature and excluding such with features with less universality. The specific classification into “more universal” vs. “less universal” features is based on the seminal work of Kroeber (1909), and is as follows:

- (14) a. More universal features
- sex of relative
 - sex of speaker
 - generation of relative
 - genealogical distance
 - affinity of relative

- lineality
 - seniority within one generation
- b. Less universal features
- sex of 1st link
 - sex of 2nd link
 - generation of 1st link
 - generation of last link
 - affinity of 1st link
 - affinity of last link

There are two grounds for this criterion. First, from a purely methodological perspective, it would be preferable to use features that are universally present in languages rather than such that are more language-specific. This is so because, in cases in which we are not forced, we should not unnecessarily extend analytical repertoire (Occam's razor). And, secondly, from a cognitive perspective, the more universal features in the left column look more salient than the less universal. To take just one example, the feature "sex." marking the sex of the referenced relative is more salient than the feature "sex 2nd link." marking the sex of just one linking relative in the chain of potentially many relatives forming a kin type.

The application of the second simplicity criterion amounts to finding: (i) the obligatory dimensions (occurring in each dimension set) and the optional dimensions (all others), and (ii) choosing, from the set of dimension sets of equal size, those sets that contain the least number of less universal features. Regarding step (i), the obligatory dimensions are "sex." "sex 1st link." "sex 2nd link." and "distance." and the optional dimensions are all remaining ones. Regarding step (ii), consider Table 4, which lists the seven equal-sized dimension sets where optional less universal dimensions are highlighted in italics.

Table 4. Dimension Sets with the Optional Less Universal Dimensions Highlighted (in Italics)

| No | Feature (dimension) sets | No |
|-----|--|----------|
| I | {sex 2nd link, sex 1st link, sex, <i>generation last link</i> , <i>generation 1st link</i> , distance, <i>affinity 1st link</i> } | 3 |
| II | {sex 2nd link, sex 1st link, sex, <i>generation last link</i> , generation, distance, <i>affinity 1st link</i> } | 2 |
| III | {sex 2nd link, sex 1st link, sex, generation, affinity, distance, <i>affinity 1st link</i> } | 1 |
| IV | {sex 2nd link, sex 1st link, sex, <i>generation last link</i> , distance, <i>generation 1st link</i> , <i>affinity last link</i> } | 3 |
| V | {sex 2nd link, sex 1st link, sex, <i>generation 1st link</i> , distance, generation, <i>affinity last link</i> } | 2 |
| VI | {sex 2nd link, sex 1st link, sex, <i>generation last link</i> , distance, generation, <i>affinity last link</i> } | 2 |
| VII | {sex 2nd link, sex 1st link, sex, affinity, distance, generation, <i>affinity last link</i> } | 1 |

In Table 4 we observe that the dimension sets with the minimum number of optional less universal features are Set III = {sex 2nd link, sex 1st link, sex, generation, affinity, distance, *affinity 1st link*} and Set VII = {sex 2nd link, sex 1st link, sex, affinity, distance, generation, *affinity last link*}, containing just one less universal feature. Accordingly, we choose these two dimension sets as simplest and discard all other sets.

7.2. Constraints on Kin Term Definitions

We can now illustrate briefly how the two constraints on kin term definitions work. The componential scheme corresponding to the simplest dimension set III was listed as Table 3 in the previous section, and we will use it for our illustrative purposes.

We start with the third simplicity criterion, preferring shorter kin

term definitions to longer ones. For instance, the word *brataničina* ‘brother’s daughter’ in Table 3 is defined in two alternative ways:

(15) *brataničina* ‘brother’s daughter’

| Alternatives | Feature sets | Size |
|--------------|---|------------|
| 1 | [sex 1st link=m, sex=f, generation= -1, affinity=cons] | 4 features |
| 2 | [sex 1st link=m, sex 2nd link=f, generation= -1, distance=2, affinity=cons] | 5 features |

Our third criterion chooses the first definition comprising 4 features, discarding the second, having 5 features. The constraint is even more effective in cases like the word *ujak* ‘mother’s brother,’ where from 5 alternatives the criterion isolates a single simplest one (the fifth one, having 4 features) and discards the other four (having 5 features):

(16) *ujak* ‘mother’s brother’

| Alternatives | Feature sets | Size |
|--------------|--|------------|
| 1 | [sex 1st link=f, affinity 1st link=cons, sex 2nd link=m, sex=m, generation=1] | 5 features |
| 2 | [sex 1st link=f, affinity 1st link=cons, sex 2nd link=m, generation=1, distance=2] | 5 features |
| 3 | [sex 1st link=f, affinity 1st link=cons, sex=m, generation=1, distance=2] | 5 features |
| 4 | [sex 1st link=f, sex 2nd link=m, generation=1, distance=2, affinity=cons] | 5 features |
| 5 | [sex 1st link=f, sex=m, generation=1, affinity=cons] | 4 features |

Even after the application of the third criterion, there still remain

many terms with multiple definitions and to these we apply the fourth simplicity criterion, which is computed as follows. Let us consider the word *nećak* ‘brother’s son or sister’s son’ as an example.

(17) *nećak* ‘brother’s son or sister’s son’

| Alternatives | Feature sets |
|--------------|---|
| 1 | [sex 2nd link=m, generation= -1, distance=2, affinity=cons] |
| 2 | [sex=m, generation= -1, distance=2, affinity=cons] |

From the set of alternative definitions of each term, we compute their intersection, which gives the *obligatory* features that must be used, whatever alternative definition for that term is chosen; the remaining features are *optional*. For example, in the case of *nećak* ‘brother’s son or sister’s son’ above, the intersection, or obligatory features, are “generation.” “distance.” and “affinity.” as they recur in the two equally simple definitions, whereas the features “sex 2nd link” and “sex” are optional and may or may not be used depending on what definition of the term we choose. We then compute the frequencies of obligatory uses of features for all terms comprising our data set. The frequency hierarchy thus obtained reflects the salience of features, the more frequent features being more salient than the less frequent ones. The computed frequency hierarchy in our case is as follows:

(18) The frequency hierarchy of obligatory features
 generation (44) > distance (27) > sex 1st link (22) >
 affinity (18) > sex (14) > affinity 1st link (9) > sex 2nd
 link (6)

Given this frequency hierarchy, distinguishing more salient from

less salient features, our criterion prefers definitions consisting of more frequent, focal, rather than less frequent, peripheral, features. In our illustrative case of *nećak* ‘brother’s son or sister’s son,’ where the two equally simple definitions differ only with respect to the use of the optional features “sex 2nd link” and “sex.” we prefer the latter having frequency=14 to “sex 2nd link” having frequency=6. We thus end up with a single definition of *nećak* ‘brother’s son or sister’s son,’ the second definition [sex=m, generation=-1, distance=2, affinity=cons]. In cases when two definitions differ in more than one feature, we select the one whose sum of optional feature frequencies is larger.

The fourth simplicity criterion is intuitive and principled as the other three criteria. An important consequence of its application is the insurance of coherence among the definitions of different kin terms in one componential model. In particular, it avoids undesired “cross-definitions” in a componential model. Thus, semantically similar terms like *nećak* ‘brother’s son or sister’s son’ and *nećakinja* ‘brother’s daughter or sister’s daughter,’ with definitions (19) and (20), will be defined with the same features and will differ only as regards the value (male or female) of one of these features, viz. sex.

(19) *nećak* ‘brother’s son or sister’s son’

| Alternatives | Feature sets |
|--------------|---|
| 1 | [sex 2nd link=m, generation= -1, distance=2, affinity=cons] |
| 2 | [sex=m, generation= -1, distance=2, affinity=cons] |

(20) *nećakinja* ‘brother’s daughter or sister’s daughter’

| Alternatives | Feature sets |
|--------------|---|
| 1 | [sex 2nd link=f, generation= -1, distance=2, affinity=cons] |
| 2 | [sex=f, generation= -1, distance=2, affinity=cons] |

Our criterion will force the choices 2 in both *nećak* and *nećakinja*, but will avoid the “cross-definition” involving choice 1 in *nećak*, but choice 2 in *nećakinja* (or vice versa). This coordination, or cohesion, between the definitions of terms in the whole kinship domain corresponds to the intuitions of native speakers. Thus, e.g., Romney & D’Andrade (1964: 156) have shown that semantically similar words like English *father* and *mother*, differing only as regards the values of a single feature, are placed by subjects adjacently in 98 per cent of the time in a free recall experiments.

The application of the two constraints on kin term definitions removes all indeterminacy and we obtain a unique componential model corresponding to the componential scheme in Table 3 with the dimension set {sex 2nd link, sex 1st link, sex, generation, affinity, distance, affinity 1st link}. This simplest model of Serbo-Croatian is listed as Table 5 below.

Table 5. Simplest Componential Model of Serbo-Croatian with the Dimension Set {sex 2nd link, sex 1st link, sex, generation, affinity, distance, affinity 1st link}

| | |
|----------------|-----------------------------------|
| <i>praded</i> | [sex=m, generation=3] |
| <i>prababa</i> | [sex=f, generation=3] |
| <i>djed</i> | [sex=m, generation=2] |
| <i>baba</i> | [sex=f, generation=2] |
| <i>otac</i> | [sex=m, generation=1, distance=1] |
| <i>majka</i> | [sex=f, generation=1, distance=1] |

| | |
|-------------------------|--|
| <i>stric</i> | [sex 1st link=m, sex=m, generation=1, distance=2, affinity=cons] |
| <i>strina</i> | [sex 1st link=m, sex=f, generation=1, distance=3] |
| <i>ujak</i> | [sex 1st link=f, sex=m, generation=1, affinity=cons] |
| <i>ujna</i> | [sex 1st link=f, sex=f, generation=1, distance=3] |
| <i>tetka</i> | [sex=f, generation=1, distance=2, affinity=cons] |
| <i>tetak</i> | [sex=m, generation=1, distance=3] |
| <i>brat</i> | [sex=m, generation=0, distance=1, affinity=cons] |
| <i>sestra</i> | [sex=f, generation=0, distance=1, affinity=cons] |
| <i>brat od strica</i> | [sex 1st link=m, sex 2nd link=m, sex=m, generation=0, distance=3, affinity=cons] |
| <i>brat od ujaka</i> | [sex 1st link=f, sex 2nd link=m, sex=m, generation=0, affinity=cons] |
| <i>brat od tetke</i> | [sex 2nd link=f, sex=m, generation=0, distance=3, affinity=cons] |
| <i>sestra od strica</i> | [sex 1st link=m, sex 2nd link=m, sex=f, generation=0, affinity=cons] |
| <i>sestra od ujaka</i> | [sex 1st link=f, sex 2nd link=m, sex=f, generation=0, distance=3, affinity=cons] |
| <i>sestra od tetke</i> | [sex 2nd link=f, sex=f, generation=0, distance=3, affinity=cons] |
| <i>bratić</i> | [sex=m, generation=0, distance=3, affinity=cons] |
| <i>sestrična</i> | [sex=f, generation=0, distance=3, affinity=cons] |
| <i>sin</i> | [sex=m, generation= -1, distance=1] |
| <i>ćerka</i> | [sex=f, generation= -1, distance=1] |
| <i>nećak</i> | [sex=m, generation= -1, distance=2, affinity=cons] |
| <i>nećakinja</i> | [sex=f, generation= -1, distance=2, affinity=cons] |
| <i>bratanić</i> | [sex 1st link=m, sex=m, generation= -1, distance=2] |
| <i>brataničina</i> | [sex 1st link=m, sex=f, generation= -1, affinity=cons] |
| <i>sestrić</i> | [sex 1st link=f, sex=m, generation= -1, affinity=cons] |
| <i>sestričina</i> | [sex 1st link=f, sex=f, generation= -1, distance=2] |
| <i>unuk</i> | [sex=m, generation= -2, affinity=cons] |
| <i>unuka</i> | [sex=f, generation= -2, affinity=cons] |
| <i>praunuc</i> | [sex=m, generation= -3, distance=3] |
| <i>praunuka</i> | [sex=f, generation= -3, distance=3] |
| <i>maž</i> | [sex=m, distance=1, affinity=aff] |

| | |
|-------------------|--|
| <i>žena</i> | [sex=f, distance=1, affinity=aff] |
| <i>svekar</i> | [sex 1st link=m, sex=m, generation=1, distance=2, affinity=aff] |
| <i>svekarva</i> | [sex 1st link=m, sex=f, generation=1, distance=2, affinity=aff] |
| <i>punac</i> | [sex 1st link=f, sex=m, generation=1, distance=2, affinity=aff] |
| <i>punica</i> | [sex 1st link=f, sex=f, generation=1, distance=2, affinity=aff] |
| <i>zet</i> | [affinity 1st link=cons, sex=m, distance=2, affinity=aff] |
| <i>snaha</i> | [affinity 1st link=cons, sex=f, distance=2, affinity=aff] |
| <i>šurjak</i> | [sex 1st link=f, affinity 1st link=aff, sex=m, generation=0, distance=2] |
| <i>šurjakinja</i> | [sex 1st link=f, affinity 1st link=aff, sex=f, distance=3] |
| <i>šogorica</i> | [sex 1st link=f, sex=f, generation=0, distance=2] |
| <i>dever</i> | [sex 1st link=m, sex=m, generation=0, distance=2] |
| <i>badženjak</i> | [affinity 1st link=aff, sex=m, distance=3] |
| <i>jetrva</i> | [sex 1st link=m, affinity 1st link=aff, distance=3] |
| <i>zaova</i> | [sex 1st link=m, affinity 1st link=aff, sex=f, generation=0, distance=2] |
| <i>prija</i> | [affinity 1st link=cons, sex=f, generation=0, distance=3, affinity=aff] |
| <i>prijatelj</i> | [affinity 1st link=cons, sex=m, generation=0, distance=3, affinity=aff] |

We recollect that our simplicity constraints on dimension sets left us with yet another simplest dimension set of 7 features, viz. {sex 2nd link, sex 1st link, sex, affinity, distance, generation, affinity last link}, differing from the first set only in the inclusion of the feature “affinity last link” in place of “affinity 1st link.” The application of the simplicity restrictions on kin term definition, again, results in the complete elimination of the kin terms indefiniteness. The second unique simplest componential model of Serbo-Croatian is listed below as Table 6.

Table 6. Another Simplest Componential Model of Serbo-Croatian with the Dimension Set {sex 2nd link, sex 1st link, sex, affinity, distance, generation, affinity last link}

| | |
|-------------------------|--|
| <i>praded</i> | [sex=m, generation=3] |
| <i>prababa</i> | [sex=f, generation=3] |
| <i>djed</i> | [sex=m, generation=2] |
| <i>baba</i> | [sex=f, generation=2] |
| <i>otac</i> | [sex=m, generation=1, distance=1] |
| <i>majka</i> | [sex=f, generation=1, distance=1] |
| <i>stric</i> | [sex 1st link=m, sex=m, generation=1, distance=2, affinity=cons] |
| <i>strina</i> | [sex 1st link=m, sex=f, generation=1, distance=3] |
| <i>ujak</i> | [sex 1st link=f, sex=m, generation=1, affinity=cons] |
| <i>ujna</i> | [sex 1st link=f, sex=f, generation=1, distance=3] |
| <i>tetka</i> | [sex=f, generation=1, distance=2, affinity=cons] |
| <i>tetak</i> | [sex=m, generation=1, distance=3] |
| <i>brat</i> | [sex=m, generation=0, distance=1, affinity=cons] |
| <i>sestra</i> | [sex=f, generation=0, distance=1, affinity=cons] |
| <i>brat od strica</i> | [sex 1st link=m, sex 2nd link=m, sex=m, generation=0, distance=3, affinity=cons] |
| <i>brat od ujaka</i> | [sex 1st link=f, sex 2nd link=m, sex=m, generation=0, affinity=cons] |
| <i>brat od tetke</i> | [sex 2nd link=f, sex=m, generation=0, distance=3, affinity=cons] |
| <i>sestra od strica</i> | [sex 1st link=m, sex 2nd link=m, sex=f, generation=0, affinity=cons] |
| <i>sestra od ujaka</i> | [sex 1st link=f, sex 2nd link=m, sex=f, generation=0, distance=3, affinity=cons] |
| <i>sestra od tetke</i> | [sex 2nd link=f, sex=f, generation=0, distance=3, affinity=cons] |
| <i>bratić</i> | [sex=m, generation=0, distance=3, affinity=cons] |
| <i>sestrična</i> | [sex=f, generation=0, distance=3, affinity=cons] |
| <i>sin</i> | [sex=m, generation= -1, distance=1] |
| <i>ćerka</i> | [sex=f, generation= -1, distance=1] |
| <i>nećak</i> | [sex=m, generation= -1, distance=2, affinity=cons] |
| <i>nećakinja</i> | [sex=f, generation= -1, distance=2, affinity=cons] |
| <i>bratanić</i> | [sex 1st link=m, sex=m, generation= -1, distance=2] |

| | |
|--------------------|--|
| <i>brataničina</i> | [sex 1st link=m, sex=f, generation= -1, affinity=cons] |
| <i>sestrić</i> | [sex 1st link=f, sex=m, generation= -1, affinity=cons] |
| <i>sestričina</i> | [sex 1st link=f, sex=f, generation= -1, distance=2] |
| <i>unuk</i> | [sex=m, generation= -2, affinity=cons] |
| <i>unuka</i> | [sex=f, generation= -2, affinity=cons] |
| <i>praunuc</i> | [sex=m, generation= -3, distance=3] |
| <i>praunuka</i> | [sex=f, generation= -3, distance=3] |
| <i>maž</i> | [sex=m, distance=1, affinity=aff] |
| <i>žena</i> | [sex=f, distance=1, affinity=aff] |
| <i>svekar</i> | [sex 1st link=m, sex=m, generation=1, distance=2, affinity=aff] |
| <i>svekarva</i> | [sex 1st link=m, sex=f, generation=1, distance=2, affinity=aff] |
| <i>punac</i> | [sex 1st link=f, sex=m, generation=1, distance=2, affinity=aff] |
| <i>punica</i> | [sex 1st link=f, sex=f, generation=1, distance=2, affinity=aff] |
| <i>zet</i> | [sex=m, affinity last link=aff, distance=2] |
| <i>snaha</i> | [sex=f, affinity last link=aff, distance=2] |
| <i>šurjak</i> | [sex 1st link=f, sex=m, affinity last link=cons, generation=0, distance=2] |
| <i>šurjakinja</i> | [sex 1st link=f, sex=f, affinity last link=aff, generation=0, distance=3] |
| <i>šogorica</i> | [sex 1st link=f, sex=f, generation=0, distance=2] |
| <i>dever</i> | [sex 1st link=m, sex=m, generation=0, distance=2] |
| <i>badženjak</i> | [sex=m, affinity last link=aff, generation=0, distance=3] |
| <i>jetrva</i> | [sex 1st link=m, affinity last link=aff, generation=0, distance=3] |
| <i>zaova</i> | [sex 1st link=m, sex=f, affinity last link=cons, generation=0, distance=2] |
| <i>prija</i> | [sex=f, affinity last link=cons, distance=3, affinity=aff] |
| <i>prijatelj</i> | [sex=m, affinity last link=cons, distance=3, affinity=aff] |

We end up with just two componential models of Serbo-Croatian. Thus, out of an astronomical number of

theoretically possible models, our simplicity constraints successfully select just two intuitive simplest models. These two models employ the dimension sets {sex 2nd link, sex 1st link, sex, generation, affinity, distance, affinity 1st link} (Table 5) and {sex 2nd link, sex 1st link, sex, affinity, distance, generation, affinity last link} (Table 6).

These models reflect the semantic structure of the Serbo-Croatian kinship domain. Each kin term is described with guaranteed necessary and sufficient components, so that leaving out a component will fail a demarcation between some kin terms, while adding a component will lead to redundancy in some kin term definition. Also, the sets of features used are guaranteed-simplest, satisfying the common linguistic requirement for parsimony.

8. Conclusions

In the paper, we showed that despite its long traditions, the practicing of the method of componential analysis encounters two major problems. These pertain to ensuring necessary and sufficient definitions of kin terms and simplicity constraints to reduce the formidable indeterminacy of componential analyses resulting from analyzing one dataset. We introduced a computer program that can nicely handle both these problems and illustrated its performance on the previously unanalyzed and elaborate Serbo-Croatian kinship terminology. The program produced just two simplest componential models out of an astronomical number of theoretically possible ones.

We have successfully analyzed more than twenty languages of various types with our program, but a detailed comparison of Serbo-Croatian with these languages is beyond the scope of the paper and we will confine to one remark. The Serbo-Croatian kin

terms form an elaborate system of 51 terms of Sudanese type, in which there are separate terms for almost each one of ego's kin, based on their distance from ego, their relation, and their gender. The Serbo-Croatian system is generally more complex than that of some other Slavic languages. Thus, e.g., Bulgarian has 43 kin terms, Polish 35, and Czech 23. The number of dimensions needed to describe the kinship domain in these languages is smaller, and is respectively 6, 6, and 4. As could be expected for genealogically related languages, they use basically the same dimensions. Only Serbo-Croatian and Polish employ the dimension "sex 2nd link." which is relatively rare cross-linguistically (but appears e.g., in Hindi and Turkish, the latter being a prototypical case of a Sudanese kinship system).

References

- Bernard, H. 2011. *Research Methods in Anthropology*. Walnut Creek, CA: Altamira Press.
- Burling, R. 1964. Cognition and Componential Analysis: God's Truth or Hocus-Pocus? *American Anthropologist* 66, 20-8.
- Geeraerts, D. 2010. *Theories of Lexical Semantics*. Oxford: Oxford University Press.
- Goodenough, W. 1956. Componential Analysis and the Study of Meaning. *Language* 32, 195-216.
- _____. 1965. Yankee Kinship Terminology: A Problem in Componential Analysis. *American Anthropologist* 67.5, 259-287.
- _____. 1967. Componential Analysis. *Science* 156, 1203-1209.
- Hammel, E. (ed.) 1965. *Formal Semantic Analysis*. Menasha, WI: American Anthropological Association.
- Hymes, D. 1964. Discussion of Burling's Paper ('Cognition and Componential Analysis: God's Truth or Hocus-Pocus?').

- American Anthropologist* 66, 116-119.
- Kroeber, A. 1909. Classificatory Systems of Relationship. *Journal of the Royal Anthropological Institute* 39, 77-84.
- Leech, G. 1974. *Semantics*. Harmondsworth: Pelican.
- Lounsbury, F. 1956. A Semantic Analysis of the Pawnee Kinship Usage. *Language* 32, 158-194.
- _____. 1964. The Structural Analysis of Kinship Semantics. In H. Lunt (ed.), *Proceedings of the 9th International Congress of Linguists* 1073-1090. The Hague: Mouton.
- Murdock, G. 1949. *Social Structure*. New York: Macmillan.
- Nogle, L. 1974. *Method and Theory in Semantics and Cognition of Kinship Terminology*. The Hague & Paris: Mouton.
- Pericliev, V. 2013. *Componential Analysis of Kinship Terminology: A Computational Perspective*. Basingstoke & New York: Palgrave MacMillan.
- Romney, A. & R. D'Andrade. 1964. Cognitive Aspects of English Kin Terms. *American Anthropologist* 66, 146-170.
- Valdes-Perez, R., F. Pereira, & V. Pericliev. 2000. Concise, Intelligible, and Approximate Profiling of Multiple Classes. *International Journal of Human-Computer Studies* 53, 411-436.
- Wallace, A. & J. Atkins. 1960. The Meaning of Kinship Terms. *American Anthropologist* 62, 58-80.
- Wordick, F. 1973. Another View of American Kinship. *American Anthropologist* 31, 1634-1656.