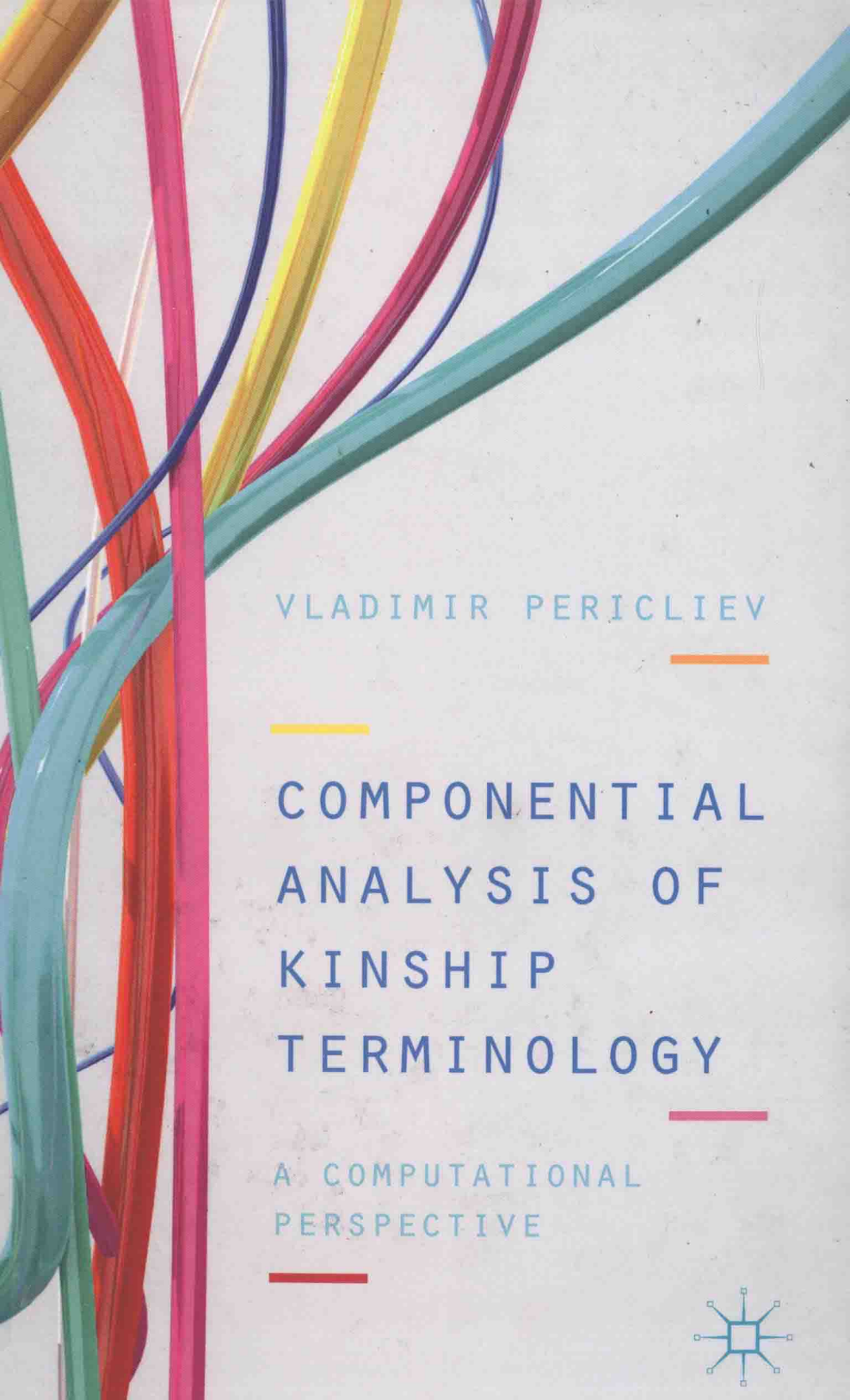

VLADIMIR PERICLIEV

COMPONENTIAL ANALYSIS OF KINSHIP TERMINOLOGY

A COMPUTATIONAL PERSPECTIVE

Componential Analysis of Kinship Terminology

A Computational Perspective

Vladimir Pericliev
Bulgarian Academy of Sciences, Sofia

palgrave
macmillan

First published 2013 by
PALGRAVE MACMILLAN

Palgrave Macmillan in the UK is an imprint of Macmillan Publishers Limited, registered in England, company number 785998, of Houndmills, Basingstoke, Hampshire RG21 6XS.

Palgrave Macmillan in the US is a division of St Martin's Press LLC, 175 Fifth Avenue, New York, NY 10010.

Palgrave Macmillan is the global academic imprint of the above companies and has companies and representatives throughout the world.

Palgrave® and Macmillan® are registered trademarks in the United States, the United Kingdom, Europe and other countries.

ISBN 978–1–137–03117–4

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources. Logging, pulping and manufacturing processes are expected to conform to the environmental regulations of the country of origin.

A catalogue record for this book is available from the British Library.

A catalog record for this book is available from the Library of Congress.

Componential Analysis of Kinship Terminology

Also by Vladimir Pericliev

MACHINE-AIDED LINGUISTIC DISCOVERY: An Introduction and Some Examples

PROFILING LANGUAGE FAMILIES BY THEIR KIN TERM PATTERNS: A Computational Approach

Preface

The notion of 'system' is central in all post-Saussurean linguistics, and componential analysis is the method for exposing the systems of linguistic entities. In componential analysis, the meaning (or value) 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 originated in phonology, but the approach was naturally extended also to the other levels of linguistic analysis: grammar and semantics. Componential analysis of kinship terminology, which lies at the crossroads of linguistics and anthropology, is the prototype example from semantics.

The early influential work on kinship semantics by Lounsbury, Goodenough and others laid the foundations of the field 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 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. This persistence in methodology is understandable insofar as rejecting the method basically implies rejecting the fundamental idea of system in linguistics. Additionally, 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.

A major goal of the present book is to critically review previous conceptions of the method and improve on previous practice of componential analysis of kin terms by setting the problem in a computational perspective. Two basic problems are isolated in previous work: the consistency of componential models, and their indeterminacy. Regarding the first problem, I will try to show, with examples

from the literature, that not infrequently proposed models are inconsistent in that they do not provide definitions of kin terms with necessary and sufficient conditions. Thus, some definitions of kin terms fail to discriminate them from the rest of the terms (failing sufficiency), whereas others contain redundant components (failing necessity). This problem has evaded previous analysts, which is no surprise, given that componential analysis is a computationally complex task (strictly speaking, NP-complete), and there existed no computational means at the time of testing proposed models. The second problem, indeterminacy, in contrast, attracted much attention owing to Robbins Burling, who warned against the existence of multiple models of the same data set and the difficulty (or even impossibility) of choosing among equally 'valid' alternatives, this circumstance putting the method in jeopardy. Dell Hymes and others objected to this view, suggesting that further constraints from the studied culture, simplicity, etc. would normally constrain the choices, but as a result of this debate no consensus was arrived at, except that what was needed were concrete empirical tests of complete kinship vocabularies rather than simplified examples and programmatic and theoretical arguments. No such empirical evidence or explicit formulation of constraints were presented to support either view, and again, the reason was the absence of computational aids to process reliably large terminological data sets.

This book describes a computer program, called KINSHIP, developed by Dr Raúl Valdés Pérez (Carnegie Mellon University) and the present author. The program is designed to resolve the above problems and, more generally, to serve as a computational tool for conducting componential analysis of kin terms. The program accepts as input the standard linguistic data, viz. the kin terms of a language with their attendant kin types (=relatives), and can generate all consistent alternative analyses (thus providing the empirical bases for evaluating the degree of indeterminacy). Additionally, the program introduces intuitive simplicity constraints on dimensions and components in kin term definitions in order to diminish, or even eliminate, multiple solutions (thus attempting to resolve the indeterminacy problem). In the book, I apply this computational machinery to complete kinship vocabularies of particular languages. The results show, first, that completely unconstrained analyses, targeting only necessary and sufficient conditions, yield an astronomical number of componential models (in accord with Burling's warning), but second, that the introduction of our natural simplicity constraints reduces this number significantly, even to unique models (tipping the balance in the direction of the opposing camp). In effect,

the operation of the KINSHIP program on concrete data from a sizable number of languages (from Indo-European and other language families) presents strong support for the power and usefulness of the method.

Our system is a sophisticated general tool for multiple-concept discrimination and therefore can be used for componential analysis of other than the kinship domain (for instance, in phonological distinctive feature analysis). Thus, for the first time after almost a century following the introduction of the concept of system and structuralism in general, linguists are capable of handling adequately the task of componential analysis, generally recognized as creative and difficult. Revealing the distinctive features of phonological and kinship systems are tasks of well-known complexity; regarding the latter, Leech (1974) for instance writes that 'kinship analyses have a mind-teasing quality of mathematical puzzles. The only cure for bafflement is to think hard and hope that the light will dawn!'

Acknowledgments

This book would have hardly been possible without the efforts of my friend and collaborator computer scientist Raúl Valdés Pérez. It was he who implemented the first version of KINSHIP, which allowed me to implement my own versions of the system (in another programming language) and continue with the linguistic investigations reported here, for which I am much indebted to him. I also gratefully acknowledge financial support for my studies on kinship terminology over the years from the International Program of the US National Science Foundation, as well as from the Bulgarian Ministry of Education and Science. Last but not least, acknowledgement is due to the Institute for Mathematics and Informatics of the Bulgarian Academy of Sciences, where I conducted much of the reported work.

Contents

List of tables in the Appendix

1
Introduction: the historical background

This chapter looks at componential analysis of kinship terminologies from a historical perspective. The underlying ideas and the basic notions of componential analysis are introduced, as described by the pioneers of the field, and some alternative approaches to describing the semantics of kinship terms are briefly sketched.

1.1 General

Every known human language has a kinship terminological system, but different languages have different organizations of these terminological systems; hence the interest of linguists and anthropologists in studying these systems. In his pioneering book *Systems of Consanguinity and Affinity of the Human Family* Henry Lewis Morgan (1871) made extensive studies of kinship terminologies of the world languages and their reflection in the social structure of society, and this work was extended and enriched by other scholars, notably anthropologist George Peter Murdock (1949). In this tradition, the meaning of kin terms in foreign languages is represented by a primitive English term (for instance 'mother, 'father', etc.), a relative product of two or more primitive terms (for instance 'mother's father') or a collection of primitive and/or relative product terms, where each primitive term and each relative product denotes a 'kin type'. This type of notation alone, useful as it is for constructing typologies of kinship terms (such as Hawaiian, Eskimo, Crow, etc. already discovered by Morgan), poses certain difficulties to the analyst regarding the important question of what the common pieces of meaning of all the kin types are that allow them to be covered by a single kin term, or what the principles of classification are of kinship in the society the anthropologist/linguist is studying. This question is

addressed by 'componential analysis', a formal procedure developed in linguistics for other purposes, and based on the Saussurian idea of 'linguistic system'.

In his *Course in General Linguistics* (*Cours de linguistique générale*, 1916), Ferdinand de Saussure created a general linguistic theory at the heart of which lay the notion of 'linguistic system' (see the brief review in Pericliev 2010: 2–3). Language (*langue*), according to Saussure, is a system in the sense that the meaning or value (*valeur*) of all linguistic entities can only be determined by their contrasts, or distinctions, from all other entities in the same system. *'In the language itself, there are only differences'*, wrote Saussure (1996[1916]: 118; italics in original), 'A linguistic system is a series of phonetic differences matched with a series of conceptual differences' (p. 118). The basic task of linguistics, then, is to reveal the structure of linguistic systems by applying the structural method of contrasts and oppositions.

Saussure's idea of language as a system broke a long tradition in Western thought dominant from Plato on, of viewing language as just an inventory of names (whatever they stood for, ideas or things in the external world), and the goal of language science as relating these names (whether derived from the true nature of things or by convention) to ready-made ideas and things given in advance of language. Saussure, in contrast to this view, conceived language not as a mere collection of discrete items, but as a highly organized totality (or, a *Gestalt*), in which the items are interrelated and derive their meaning from the system as a whole. Thus, he writes:

> In all these cases, what we find, instead of *ideas* given in advance, are *values* emanating from the linguistic system. If we say that these values correspond to certain concepts, it must be understood that the concepts in question are purely differential. That is to say they are concepts defined not positively, in terms of their content, but negatively by contrast with other items in the same system. What characterises each most exactly is being whatever the others are not. [Saussure 1996[1916]: 115; italics in original]

The importance of the idea of language as a system, and structuralism in general, cannot be overstated. It influenced researchers both in linguistics and outside of linguistics. Within linguistics, the structural method came to be recognized as an indispensable tool at all levels of linguistic analysis: phonology ('distinctive feature analysis'), semantics ('structural semantics'), morphology, etc. Transformationalists

(especially Chomsky) emphasized that formal, generative grammar as a whole is a 'systemic notion' in that a simplification in some component leads to more complexities in another component. Also, the idea gave rise to different linguistic trends like the Prague school (Trubetzkoy, Jakobson), the Copenhagen school (Hjelmslev) and American structuralism (Bloomfield, Bloch, Harris, etc.). Outside of linguistics, the principles and methods of structuralism were adopted by scholars of such diverse areas as anthropology (Claude Lévi-Strauss), psychoanalysis (Lacan) and literary criticism (Barthes) and were implemented in their respective areas of study. According to Assiter (1984), there are four common ideas regarding structuralism that form an 'intellectual trend'. First, the structure is what determines the position of each element of a whole. Second, structuralists believe that every system has a structure. Third, structuralists are interested in 'structural' laws that deal with coexistence rather than changes. And, finally, structures are the 'real things' that lie beneath the surface or the appearance of meaning.

The Saussurian idea of system, and the related formal procedures developed in linguistics for discovering oppositions in phonological and semantic systems, were transferred by direct analogy to componential analysis of kinship terminologies. The pioneers in the field, Floyd Lounsbury and Ward Goodenough, readily acknowledge this. Thus, in a seminal article in *Language*, Lounsbury states that 'The aim of this paper is to point out a relatively simple problem in semantics which can be analysed by means of techniques analogous to those already developed in linguistics [...]' (Lounsbury 1956: 158–9). Goodenough, analogously, refers to the utility and rigour of the procedures already developed in linguistics:

> Inspiration [...] has come largely from accomplishments of linguistic science. Linguists are able to produce elegant and accurate representations of what one has to know in phonology and grammar if one is to speak particular languages acceptably by native standards. Their procedures enable them to replicate one another's work readily. Application of the basic strategies of descriptive linguistics to the problem of describing other facets of culture is helping to raise the standards of rigor in ethnographic description. These strategies include what is best described as contrastive analysis. Its use for describing how people classify phenomena, insofar as their classifications are reflected in the vocabulary of their language, has led to the analytic method described here. (Goodenough 1967: 1203)

The following section describes the method of componential analysis of kinship vocabularies in more detail.

1.2 Componential analysis of kinship terminological systems: the basic notions

The method of componential analysis was introduced into kinship semantics basically through the work of anthropologists Lounsbury and Goodenough (Goodenough 1956, 1964, 1967; Lounsbury 1956, 1964, 1965; but see also Greenberg 1949; Wallace and Atkins 1960; Hammel 1965; Leech 1974, and more recently, Geeraerts 2010; Bernard 2011).

1.2.1 Kin terms and kin types

The *kin terms* of a language, such as English *mother, aunt, son-in-law*, etc., 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'.

(Another common notation for the atomic relationships is: F = father, M = mother, B = brother, Z = sister, S = son, D = daughter, H = husband, W = wife.)

Additional symbols may be used to specify relative age or sex of the speaker, for instance:

y = younger, e = elder; m = male ego, f = female ego

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', MoBre 'mother's elder brother', etc.

The meaning of kin terms is represented by all kin types, or relatives, covered by the term. For example, the (simplified) meaning of the English term *uncle* is FaBr and MoBr. The set of all kin terms in a language is the *kinship vocabulary* of the language.