ELSEVIER Evolution and Human Behavior 24 (2003) 320–350

Evolution and Human Behavior

The generative psychology of kinship Part 2. Generating variation from universal building blocks with Optimality Theory

Doug Jones*

Department of Anthropology, 102 Stewart Building, University of Utah, Salt Lake City, UT 84112, USA

Received 12 August 2002; received in revised form 10 June 2003

Abstract

Optimality Theory (OT), a new approach to rules of language, can be used to analyze variation and universals in kin classification. According to the theory, rules of language take the form of ranked constraints that filter random variation; grammatical differences between languages reflect different constraint rankings. Applied to kin terminology, OT provides an elegant account of how universal schemas of sociality generate different terminologies for aunts, uncles, and cousins that merge and separate various kin types, and why many logically possible terminologies are rare or nonexistent. The theory may help to narrow gaps between cognitive–linguistic accounts of kin terminology and sociological accounts, and between domain-specific and domain-general views of cognition. A companion paper discusses principles of kin classification and the evolutionary psychology of kinship.

Keywords: Evolutionary psychology; Human kinship; Kin selection; Kin terminology; Optimality Theory

---

1. Introduction

Anthropologists have often noted parallels between rules of language and rules of kinship and kin classification (Goodenough, 1970; Greenberg, 1966, 1990; Kroeber, 1909;

---

* 260 Greenstone Drive, Reno, NV 89512, USA.
E-mail address: douglas.jones@anthro.utah.edu (D. Jones).

doi:10.1016/S1090-5138(03)00037-0

Lévi-Strauss, 1967, 1969; Woolford, 1984), and several schools of anthropology, including structuralism and componential analysis, have argued that theories of the grammar of language are a key to understanding the "grammar" of kinship (and other domains of thought). In a companion paper, I give an informal sketch of a universal grammar of kin terminology built around trade-offs between conserving information about binary distinctive features and avoiding marked categories. This paper gives a formal exposition, using a new approach to rules of language called Optimality Theory (OT). Although proposed only recently (Prince and Smolensky, 1993), OT has already been extremely influential, inspiring work in many areas of phonology and syntax; this paper proposes extending the theory to a semantic domain. (A note on terminology: OT in linguistics is distinct from optimality theory in behavioral ecology.)

Section 2 briefly reviews OT in linguistics, showing parallels with principles of kin classification. Succeeding sections apply the theory to terminologies for aunts, uncles, and cousins, substantiating and elaborating the claims of the companion paper about cognitive universals underlying and supporting variation in kinship systems. (Applications to other kin categories, including siblings and grandkin, are the subject of further papers, including Jones, in preparation.) A final section takes stock of the advantages of optimality theory in the study of kin classification, and considers some implications for arguments about domain-specificity and domain-generality in evolutionary and cognitive psychology.

2. OT from linguistics to kinship

Knowing a language means knowing how to go from input to output. For example, speakers may need to go from a concept to a word, from the singular of a noun to its plural, from the present tense of a verb to its past tense, or from a statement to a corresponding question. In each of these cases, the move from input to output is usually highly deterministic — a speaker who cannot generate the uniquely appropriate output for a given input does not know (part of) the language.

Sometimes, speakers master input–output connections through rote memorization. With most of the lexicon, for example, there is no other way to know which word is attached to which concept. This holds also for many irregular plurals and past tenses. However, this kind of memorization cannot possibly account for all linguistic knowledge. Speakers of a language can agree on regular plurals and past tenses of words they have never heard before, and they can generate millions of grammatically correct sentences over a lifetime that no one ever produced before. Linguistic knowledge in these cases must be knowledge of a manageable number of rules that can generate appropriate outputs for an astronomical number of potential inputs (Pinker, 1999). One of the great tasks facing linguists, therefore, is to figure out the rules that speakers must know to use a language. And linguists face a further challenge: They also have the job of figuring out how these rules are learned. Very little of this knowledge is passed on through explicit instruction. How is it then that children born ignorant of any particular language infer the abstract rules underlying the utterances they hear?

2.1. From input to output through variation and selection

OT offers one powerful set of answers to these questions (Archangeli 1997; Boersma, Dekkers, & van de Weijer, 2000; Prince & Smolensky, 1993, 1997; Tesar, Grimshaw, & Prince, 1999). It differs from most previous transformational or derivational approaches, in which input–output transformations operate in a serial or assembly-line fashion. In these approaches, an input is transformed by a rule that produces a deterministic intermediate output, which is then passed on to the next rule in line, and so on, until the final output emerges. In OT, by contrast, linguistic regularities are the outcome of a process of intracranial variation and selection. The theory cashes in on the argument that cognition involves internal trial-and-error processes analogous to Darwinian natural selection (Cziko, 1995; Dennett, 1995). In OT, rules (or constraints) do not actively transform input but act as filters on random variation (Fig. 1). The theory posits three stages in the production of an output. In the first, Input stage, any linguistically well-formed input is allowed. In the second, Random Variation stage, the input gives rise to a large number of candidate outputs. In this stage, too, almost anything goes. Candidate outputs are not required to bear any resemblance to inputs, as long as they meet some broad standards of linguistic well-formedness. Even if in practice there are some mental limitations on how many candidates can be considered, such limitations are not required or assumed by OT. Finally, in the third, Selection stage, the real work of choosing an output is done. Candidates are evaluated in parallel — in effect, run through a succession of filters — until all but one is eliminated. This *optimal* candidate is the output. It is in the Selection stage that the rules of grammar — of phonology, morphology, or syntax — operate, in the form of violable or "soft" constraints. The preferred output is not that which follows all the rules: This is often impossible, because rules are often mutually inconsistent. Instead, the preferred output is that which produces the *least serious* violations of a ranked set of rules. A candidate which

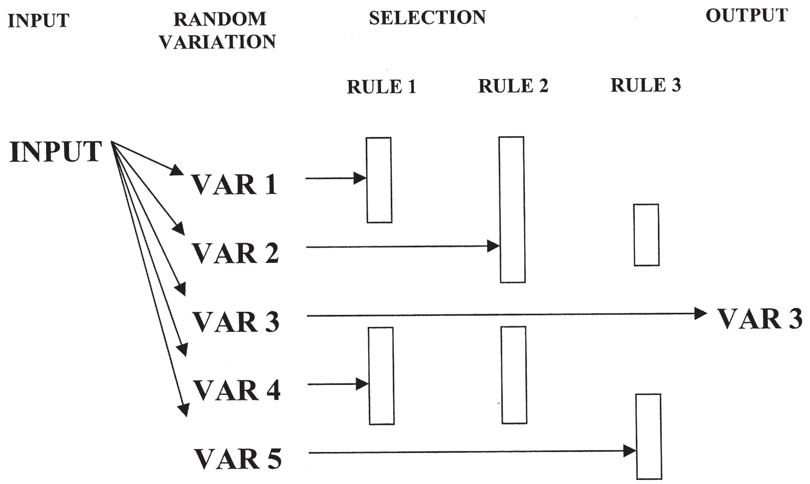

Fig. 1. OT: variation and selection.

obeys a high ranking rule will always beat a candidate which violates that rule, no matter how many more lower ranking rules the first candidate violates.

Kin terminology, like other linguistic domains in which OT has been successful, involves regular input–output relationships. An input can be represented by some formula of the form "father's brother's daughter" or "man's older sister," and the output is some native word or else a null output if the corresponding word does not exist in the language in question. Of course, these input–output relationships are partly a matter of rote memorization of arbitrary words. But even if a generative approach like OT cannot hope to predict the specific words used for kin, it might still predict typological regularities — whether two different inputs will yield the same or different outputs. This implies that when applied to a semantic field like kinship, OT may involve a different balance between rules and memorization than when applied to grammar. In the case of grammar, rules are called on constantly to produce and interpret novel utterances. In the case of kin terminology, however, generative principles may be active not so much on a daily basis — when memorization of kin equations may suffice — as on the longer time scales associated with linguistic evolution. It is when meanings shift over time that the principles of OT may dictate that some shifts are much likelier than others. Although there may be a residue of irregularity in kin terminology that no system of rules can account for, we will see below that psychological constraints on the cultural evolution of kin terminology have resulted in highly regular patterns.

2.2. Faithfulness and markedness

In OT, there are two types of constraints, which can be summed up by the slogans "Conserve information" and "Keep it simple." The first type is an Input–Output or faithfulness constraint. It states that the output should be similar to the input along some dimension. For example, when English dialects have dropped the initial consonant sounds from loanwords like "tsar" and "psyche," they have violated a faithfulness constraint that requires input to match up with output.

The second type of constraint is an Output or markedness constraint. It states a preference for a particular kind of output, regardless of what the input is. The concept of markedness is discussed by Prince and Smolensky (1997):

> [A]n element of linguistic structure is said to be marked if it is more complex than an alternative along some dimension; the relevant dimensions may sometimes correlate with comprehension, production, memory, or related physical and cognitive functions. ... Marked elements tend to be absent altogether in certain languages, restricted in their use in other languages, later acquired by children, and in other ways avoided.

An unmarked linguistic structure, by contrast, is structurally simple. It is, in the language of cognitive psychology, the prototype or default form. Markedness constraints in OT state that unmarked forms are preferred; marked forms violate these constraints. For example, there is a universal preference for single consonant over multiple consonant syllable onsets; the latter are marked relative to the former.

Obviously, these two types of constraints look a lot like the descriptive and classificatory tendencies in kin terminology noted by Kroeber (1909) and Morgan (1870) and discussed in the companion paper.

2.3. Universal constraints and variable rankings

OT claims to account not only for the regularities of specific languages, but for language universals as well. Roughly speaking, the theory claims that grammatical variation reflects permutations in a universal constraint set — that all languages have the same constraints, while grammatical differences between languages result from differences in the rankings given to different constraints. For example, in some languages, like Japanese, a markedness constraint banning multiconsonant syllable onsets ranks very high, and all syllables begin with single consonants. In English, the same constraint ranks somewhat lower: "star" and "spiky" faithfully retain multiconsonant onsets, but generally not "tsar" or "psyche." In Russian and Greek, the constraint ranks even lower relative to faithfulness constraints. Thus, a *No multiconsonant onset* constraint can generate a wide range of surface patterns depending on its rank. But some patterns cannot be generated by any possible ranking, such as a language with only multiconsonant syllable onsets. Empirically, this and many other logically possible languages are never found.

Several complications that will prove relevant below should be noted: In current versions of OT, not all constraints are active in all languages. Instead, there is a smaller set of universal constraint *templates* from which language-specific constraints are derived. And there are also universal constraint hierarchies or gradients of the form "constraint A always outranks constraint B" that rule out some permutations of constraint rankings.

Regular input–output relationships; distinctive features; markedness; universal principles with cross-cultural variation in the priority given to different principles; derivation of constraints and constraint hierarchies from a small set of constraint schemas — these are the stuff of linguistic OT and also, apparently, of kin classification. Below, I show how OT applies to kin classification, specifically to terminology for aunts, uncles, and first cousins. The analysis does not consider any one culture's kin terminology in detail, but it reviews typological variation across cultures, drawing particularly on Murdock's (1970) large data set on kin terminology. Just one aspect of markedness is considered, namely syncretization, although other phenomena such as marking sensu stricto and par excellence expression could also be handled with modest revisions.

3. Aunts and uncles

As discussed in the companion paper, terminologies for aunts and uncles fall into four major categories. This section begins with these, in order to introduce the basics of OT as applied to kin classification. It then demonstrates some of the flexibility of the theory with a finer-grained analysis of aunt and uncle terminologies and their interrelations. The analysis considers only terms for consanguineals, ignoring aunts and uncles by marriage.

3.1. Aunts: basic categories

In OT, the process of variation and selection leading to regular input–output relations is commonly presented in the form of one or more *tableaus*. Table 1 shows three tableaus associated with the lineal aunt terminology familiar to English speakers, in which "mother's sister" and "father's sister" are equated with one another (as "aunt") but distinguished from "mother." In the upper left-hand cell of each tableau is an input, corresponding to the Input stage in Fig. 1. The input is "mother" (M) in the first tableau, "mother's sister" (MZ) in the second, and "father's sister" (FZ) in the third. The cells in the left-hand columns below these

Table 1
Lineal aunts: tableaus

Mother (M) is "mother" (M) . . .

M	DLin	*FZ	*MZ	DBif
M✓				
MZ	*!		*	
FZ	*!	*		*

mother's sister (MZ) is "mother's sister" (MZ) . . .

MZ	DLin	*FZ	*MZ	DBif
M	*!			
MZ✓			*	
FZ		*!		*

and father's sister (FZ) is also "mother's sister" (MZ).

FZ	DLin	*FZ	*MZ	DBif
M	*!			*
MZ✓			*	
FZ		*!		*

So there is one term for mother, and another for mother's sister and Father's Sister.

Descriptive and classificatory constraints:

DLin	*Distinguish lineal and collateral kin*
*FZ	No "father's sister"
*MZ	No "mother's sister"
DBif	*Distinguish maternal and paternal kin*

Other abbreviations:

*	Constraint violation
*!	Fatal constraint violation (following cells are shaded; do not affect optimal output)
✓	Optimal output

inputs show candidate outputs, corresponding to the Random Variation stage in Fig. 1. For each tableau, we consider three candidate outputs, M, MZ, and FZ.

The row across the top of each tableau shows a ranked series of constraints that evaluate each candidate output, corresponding to the rules in the Selection stage in Fig. 1. The order of the constraints is constant throughout Table 1, and in any one system of terminology, although we will see below that it can vary across cultures. The present constraint ranking is:

1. *DLin*: *Distinguish lineal and collateral kin*
2. **FZ*: *No "father's sister"*
3. **MZ*: *No "mother's sister"*
4. *DBif*: *Distinguish maternal and paternal kin*

or, more compactly:

$$\mathit{DLin} \gg {}^{*}\mathit{FZ} \gg {}^{*}\mathit{MZ} \gg \mathit{DBif}$$

The first and last constraints, *DLin* and *DBif*, are descriptive (or Input–Output or faithfulness) constraints, which require that input and output match up with regard to particular distinctive features. Thus, *DLin* requires that if the input is a lineal relative, then the output must also be a lineal relative, and similarly for collateral relatives. The first tableau shows the operation of this constraint with the input "mother." Because parents are lineal relatives and parents' siblings are collaterals, the Input → Output transformation M → M obeys *DLin*, while M → MZ and M → FZ violate it. The result is shown in the first tableau in the column under *DLin*, where constraint violations are marked with asterisks. *DLin* is the top ranked constraint, and any candidate violating this constraint is removed from further consideration. Exclamation marks identify these fatal constraint violations, and gray shading shows cells that are out of play as a result, and can have no effect on the outcome. The only surviving candidate output, M, is marked with a check. Thus, the optimal output for "mother" is "mother."

The other descriptive constraint, *DBif*, in the last column, requires that if the input is a maternal relative, then the output must be a maternal relative, and similarly for paternal relatives. Because "mother" and "mother's sister," but not "father's sister," are maternal relatives, either M → M or M → MZ obeys *DBif*, while M → FZ violates it, as shown. However, the bifurcate distinction ranks low in the present case, so this has no effect on kin terminology.

The other two constraints, **FZ* and **MZ*, are classificatory (or Output or markedness) constraints, which ban certain "marked" outputs regardless of the input. The **FZ* constraint bans FZ as output, and the **MZ* constraint bans MZ, as shown by asterisks in the appropriate cells. Consistent with the markedness of both types of "aunt" relative to "mother," there is *no *M* constraint banning terms for mother. In the first tableau, the two classificatory constraints have no bearing on the outcome. In the second and third tableaus, however, they ensure that the optimal output for both MZ and FZ is MZ, because **FZ* outranks **MZ*, thus eliminating the MZ → FZ and FZ → FZ rows. MZ is not a perfect output, because it violates the **MZ* constraint, but it is *optimal*, because it produces the least serious violations given the constraint ranking. The equation of "father's sister" and "mother's sister" in the third tableau

(FZ → MZ) also violates the classificatory constraint *DBif*, but this violation has no effect. Thus, the final outcome of the present constraint ranking is a terminology with one term for "mother" and a second term covering both "mother's sister" and "father's sister."

Note that in the present case, the term for "mother's sister" and "father's sister" is rendered "mother's sister." However, if the two classificatory constraints **FZ* and **MZ* were switched around (i.e., *MZ $\gg$ *FZ), then "mother's sister" and "father's sister" would instead both be equated with "father's sister." The actual outputs would be different from those in Table 1, but the more fundamental result — a single term for mother and a single term for parent's sister — would be the same. Because **FZ* and **MZ* can be placed in either order without changing the typological outcome, we say that they belong to the same *stratum*. In a tableau, constraints within a stratum are separated by dotted rather than solid lines. Membership in a stratum can also be indicated by parentheses, (**FZ *MZ*), with the rank inequality sign, $\gg$, omitted. The constraint ranking for lineal aunt terminologies given above can thus be restated:

$$DLin \gg (*FZ\ *MZ) \gg DBif$$

or equivalently:

$$DLin \gg *PZ \gg DBif$$

where **PZ* (*No "parent's sister"* or, more informally, *No "aunt"*) is equivalent to the stratum (**FZ *MZ*). Psychologically, this implies that in cultures with lineal aunt terminologies, neither type of aunt is consistently more prototypical, at least insofar as cognition is given linguistic expression.

The preceding analysis demonstrates how OT generates lineal aunt terminology from a constraint ranking in which the preservation of lineal/collateral distinctions outranks the markedness of aunt terms, which, in turn, outranks the preservation of bifurcate maternal/paternal distinctions. In OT, alternative patterns can be generated by changing constraint rankings. For example, Table 2 shows what happens when *DLin* and *DBif* are switched around to give:

$$DBif \gg *PZ \gg DLin$$

In this case, in the second tableau, the constraint **MZ* requires that the input "mother's sister" be replaced by something else, although it does not dictate what the replacement will be. Because the bifurcate distinction *DBif* ranks high, the maternal MZ cannot be replaced with the paternal FZ. Instead, the optimal replacement is MZ → M. In the third tableau, the high-ranking *DBif* prevents the paternal FZ input from generating the maternal M or MZ as output. The overall result is a bifurcate merging terminology ("mother" = "mother's sister" ≠ "father's sister").

Further permutations generate the remaining major categories (tableaus omitted here). In the maximally classificatory one-term generational terminology ("mother" = "mother's

Table 2
Bifurcate merging aunts: tableaus
Mother (M) is “mother” (M) . . .

M	*DBif*	*FZ	*MZ	*DLin*
M✓				
MZ			*!	*
FZ	*!	*		*

mother’s sister (MZ) is also “mother” (MZ) . . .

MZ	*DBif*	*FZ	*MZ	*DLin*
M✓				*
MZ			*!	
FZ	*!	*		

but father’s sister (FZ) is “father’s sister” (FZ).

FZ	*DBif*	*FZ	*MZ	*DLin*
M	*!			*
MZ	*!		*	
FZ✓		*		

So there is one term for mother and mother’s sister, and another for father’s sister.

See Table 1 for definitions and abbreviations.

sister” = “father’s sister”), the imperative of avoiding marked aunt terms overrides all descriptive constraints:

$$*PZ \gg (DLin\ DBif)$$

Here, the top rank of the $*PZ \equiv (*FZ\ *MZ)$ constraint ensures that the optimal output for both MZ and FZ is M. (The ≡ sign indicates that $*PZ$ and $(*FZ\ *MZ)$ are equivalent expressions.)

At the other extreme, the maximally descriptive three-term bifurcate collateral terminology (“mother” ≠ “mother’s sister” ≠ “father’s sister”) reverses this ranking:

$$(DLin\ DBif) \gg *PZ$$

In both of these cases, the order of the two descriptive constraints is irrelevant.

These four terminologies are the only ones that can be generated by permuting the constraints above, even if we split the $*PZ$ stratum into its two components and move these around separately. As long as *DLin* and *DBif* are the only tools available to enforce distinctions, there is no way to generate the fifth logically possible system (“mother” = “father’s sister” ≠

"mother's sister"). Of course, this is a strength of the present account, because this system is found nowhere among the world's languages (Hage, 1997).

3.2. Aunts: relative age and markedness

The preceding analysis demonstrates some of the basic principles of linguistic OT applied to an important area of kin terminology. However, it does not show that OT has any advantages compared to a simple componential analysis with two distinctive features, lineal/collateral and bifurcate maternal/paternal, each with two possible settings. We can get a better idea of the potential of OT in kin classification by introducing a third distinctive feature active in some aunt terminologies, relative age.

Many languages have separate terms for aunts depending on whether they are older or younger sisters of the linking parent. Murdock (1970) provides data on aunt terminologies as part of the most thorough available review of cross-cultural variation in kin terminologies. Of the 566 cultures in his sample, 449 employ one of the four basic terminologies discussed above (151 bifurcate collateral, 143 bifurcate merging, 89 lineal, and 66 generational). Just three types of terminology employing relative age distinctions account for 93 of the remaining 117 terminologies. The *skewed bifurcate collateral* pattern, with terms for "mother," "mother's older sister," "mother's younger sister," and "father's sister," accounts for 64 cases. The *relative age* pattern ("mother," "parent's older sister," "parent's younger sister") accounts for 15 cases. The *age-differentiated bifurcate collateral* pattern ("mother," "mother's older sister," "mother's younger sister," "father's older sister," "father's younger sister") accounts for 14 cases. Of the remaining 24 cases, 13 involve sex-of-speaker distinctions, where terms differ depending on whether a man or woman is speaking. These are not analyzed further here. Another eight are miscellaneous rare types, and for nine, no data or only questionable data were available.

Table 3 provides four tableaus illustrating how OT generates the most common terminology with relative age distinctions, the skewed bifurcate collateral pattern. This pattern follows the bifurcate collateral pattern in preserving both lineal/collateral and bifurcate maternal/paternal distinctions. But it also adds an older/younger sibling distinction, applied in a skewed fashion to "mother's sister" but not to "father's sister." The relevant constraint ranking is:

$$(DBif\ \ DLin) \gg {}^{*}FZ \gg DAge \gg {}^{*}MZ$$

In common with the bifurcate collateral ranking, this assigns top rank to *DLin* and *DBif*. However, it adds a new descriptive constraint, *DAge*, which requires that if the input is an older sibling, then the output must be an older sibling of some kind, and similarly for younger siblings. The range of allowed inputs is revised to allow for relative age differences to include five kin types: "mother," "mother's older sister" "mother's younger sister," "father's older sister," and "father's younger sister." And the two classificatory constraints **FZ* and **MZ* are now expanded into constraint strata, $^{*}FZ \equiv (^{*}FyZ\ ^{*}FoZ)$ and $^{*}MZ \equiv (^{*}MyZ\ ^{*}MoZ)$.

The first two tableaus show the two types of "mother's sister" as inputs. In both cases, the first two constraints, *DBif* and *DLin*, rule out M, FoZ, and FyZ as output. The next two

Table 3

Skewed bifurcate collateral aunts: tableaus

On the one hand, mother's older sister (MoZ) is "mother's older sister" (MoZ) . . .

MoZ	*DBif*	*DLin*	**FZ ≡ (*FyZ*	**FoZ)*	*DAge*	**MZ ≡ (*MyZ*	**MoZ)*
M		*!			*		
MoZ✓							*
MyZ					*!	*	
FoZ	*!			*			
FyZ	*!		*		*		

while mother's younger sister (MyZ) is "mother's younger sister" (MyZ).

MyZ	*DBif*	*DLin*	**FZ ≡ (*FyZ*	**FoZ)*	*DAge*	**MZ ≡ (*MyZ*	**MoZ)*
M		*!			*		
MoZ					*!		*
MyZ✓						*	
FoZ	*!			*	*		
FyZ	*!		*				

On the other hand, while father's older sister (FoZ) is "father's older sister" (FoZ) . . .

FoZ	*DBif*	*DLin*	**FZ ≡ (*FyZ*	**FoZ)*	*DAge*	**MZ ≡ (*MyZ*	**MoZ)*
M	*!	*			*		
MoZ	*!						*
MyZ	*!				*	*	
FoZ✓				*			
FyZ			*!		*		

father's younger sister (FyZ) is *also* "father's older sister" (FoZ).

FyZ	*DBif*	*DLin*	**FZ ≡ (*FyZ*	**FoZ)*	*DAge*	**MZ ≡ (*MyZ*	**MoZ)*
M	*!	*			*		
MoZ	*!				*		*
MyZ	*!					*	
FoZ✓				*	*		
FyZ			*!				

So there is one term for mother (not shown here), one for mother's older sister, one for mother's younger sister, and one for father's sister, older or younger.

constraints, in the **FZ* ≡ (**FyZ* **FoZ*) stratum, are irrelevant here. The next constraint, *DAge*, rules out any merger of "mother's older sister" and "mother's younger sister," MoZ → MyZ or MyZ → MoZ, although the latter transformation would better satisfy the lower ranking **MyZ* constraint. The result is two separate terms for "mother's older sister" and "mother's younger sister."

The third and fourth tableaus, with two types of "father's sister" as input, show a different pattern. Once again, the first two constraints preserve lineal/collateral and bifurcate maternal/paternal distinctions: M, MoZ, and MyZ are ruled out as outputs. But because the markedness of "father's sister" (the **FZ* ≡ (**FyZ* **FoZ*) stratum) outranks the relative age distinction, *DAge*, the optimal output for both FoZ and FyZ is FoZ. Note that changing the order of **FyZ* and **FoZ* within the **FZ* stratum would change optimal output in both cases to FyZ without changing the more fundamental outcome: a single term for "father's sister" regardless of relative age. The overall result is a four-term terminology: "mother," "mother's older sister," "mother's younger sister," and "father's sister."

Table 4 gives constraint rankings for the seven major types of aunt terminology. For the first four types, the relative age distinction, *DAge*, ranks lower than any of the classificatory constraints, which ensures that any terms differing only in relative age will generate the same output and be terminologically equated. The relative age pattern ("mother," "parent's older sister," "parent's younger sister") results from giving a high rank to *DAge* and a low rank to the bifurcate maternal/paternal distinction, *DBif*. In the five-term age-differentiated bifurcate collateral pattern ("mother," "mother's older sister," "mother's younger sister," "father's older sister," "father's younger sister"), all three descriptive constraints rank high. The table summarizes abbreviations for kin types and constraints and constraint strata. Classificatory constraints can be expanded into strata where appropriate to input. The constraint rankings shown are the most straightforward ones that generate each terminology. In some cases, more complicated rankings with split strata would also generate the corresponding terminologies.

The distribution of aunt terms shows two important markedness effects relating to bifurcation and relative age. Markedness effects can be handled in the framework of OT with *markedness gradients*, which impose universal or near-universal limits on permitted constraint rankings. Two such gradients are given at the bottom of Table 4. The first of these, **FZ* ≥ . **MZ*, means that "father's sister" is almost universally at least as marked as "mother's sister." (Note that inequality signs for universal markedness gradients, ≥ and >, are different from the sign indicating culture-specific rankings, ≫. A period after a sign

---

Notes to Table 3:

Descriptive and classificatory constraints

DBif	*Distinguish maternal and paternal kin*
DLin	*Distinguish lineal and collateral kin*
FZ* ≡ (FyZ* **FoZ*)	*No "father's sister"* stratum equivalent to (*No "father's younger sister" No "father's older sister"*)
DAge	*Distinguish senior and junior kin (within generations) relative to parent*
MZ* ≡ (MyZ* **MoZ*)	*No "mother's sister"* stratum equivalent to (*No "mother's younger sister" No "mother's older sister"*)

Table 4
Generating aunt terms

Type	Equations	Constraint ranking
Generational	M = M = FZ	*_PZ_ ≫ (_DBif DLin DAge_)
Lineal	M MZ = FZ	_DLin_ ≫ *_PZ_ ≫ (_DBif DAge_)
Bifurcate merging	M = MZ FZ	_DBif_ ≫ *_PZ_ ≫ (_DLin DAge_)
Bifurcate collateral	M MZ FZ	(_DBif DLin_) ≫ *_PZ_ ≫ _DAge_
Relative age	M PoZ PyZ	(_DLin DAge_) ≫ *_PZ_ ≫ _DBif_
Skewed bifurcate collateral	M MoZ MyZ FZ	(_DBif DLin_) ≫* _FZ_ ≫ _DAge_ ≫ *_MZ_
Age-differentiated bifurcate collateral	M MoZ MyZ FoZ FyZ	(_DBif DLin DAge_) ≫ *_PZ_

Kin type abbreviations:	
M	Mother
Z	Sister
F	Father
oZ	Older sister
P	Parent
yZ	Younger sister
Descriptive constraints:	
DBif	_Distinguish maternal and paternal kin_
DLin	_Distinguish lineal and collateral kin_
DAge	_Distinguish senior and junior kin (within generations) relative to parent_
Classificatory constraints:	
*_FZ_ ≡ (*_FyZ_ *_FoZ_)	_No "father's sister"_ stratum equivalent to (_No "father's younger sister" No "father's older sister"_)
*_MZ_ ≡ (*_MyZ_ *_MoZ_)	_No "mother's sister"_ stratum equivalent to (_No "mother's younger sister" No "mother's older sister"_)
*_PZ_ ≡ (*_FZ_ *_MZ_)	_No "parent's sister"_ (i.e., _No "aunt"_) stratum equivalent to (_No "father's sister" No "mother's sister"_)
Markedness gradients:	
*_FZ_ ≥ . *_MZ_	Paternal (cross) aunts marked relative to maternal (parallel) aunts
*_PyZ_ = . *_PoZ_	Parent's older and younger sisters equally marked

indicates that it is a near-universal, with rare exceptions, rather than an absolute universal.) The greater-than-or-equal-to sign for this gradient means that any constraint ranking with *_FZ_ ≫ _X_ ≫ *_MZ_ is allowed, where _X_ is any other constraint. Also allowed is the constraint stratum *_PZ_ ≡ (*FZ *MZ). But any rankings of the form *_MZ_ ≫ _X_ ≫ *_FZ_ are _not_ allowed. This captures the greater markedness of "father's sister": A distinctive feature can be active in the "father's sister" category only if it is also active in the "mother's sister" category, but not conversely. In particular, as long as constraint rankings with *_MZ_ ≫ _DAge_ ≫ *_FZ_ are ruled out, a reverse-skewed bifurcate collateral pattern ("mother," "mother's sister," "father's older sister," "father's younger sister") is impossible. Although this pattern is not completely unknown, it is vastly less common than the skewed bifurcate collateral pattern: just 1 case vs. 64 cases in Murdock's sample.

The second markedness gradient active in aunt terminology, *_PyZ_ = . *_PoZ_, implies that a parent's younger and older siblings are almost universally equally marked. Thus, *_MyZ_ and *_MoZ_ must be assigned to the same *_MZ_ stratum; neither *_MyZ_ ≫ _X_ ≫ *_MoZ_ nor *_MoZ_ ≫ _X_

≫ _*MyZ_ is allowed. The same applies to _*FyZ_ and _*FoZ_. This accounts for a near-universal: When maternal/paternal bifurcation applies to "parent's older sister," it almost always applies to "parent's younger sister" as well (two exceptions), and vice versa (two exceptions).

In the companion paper, I argue that the constraints that generate kin terminologies derive from three schemas of social relations: genealogical distance, social rank, and group membership. The first markedness gradient, making "father's sister" at least as marked as "mother's sister," seems to relate to the schema of group membership. The bifurcate distinction is strongly associated with the presence of unilineal descent groups. In a society with patrilineal or matrilineal descent groups, or both, mother and mother's sister will be born into the same descent group, while father's sister will be born another descent group. (In Section 4, I take this argument further, showing that _DBif_ comes in patrilineal and matrilineal versions.) This is not to say that all societies with unilineal descent have bifurcate terminology, or vice versa, but rather that both unilineal descent and bifurcate terminology depend on the activity in the kinship domain of a schema of group membership.

In the second markedness gradient, the schema involved seems to relate to social rank. Apparently, the social ranks of parent's older and younger siblings relative to Ego are generally close enough that they are treated the same with regard to bifurcation. This makes an interesting contrast with terminology for one's own siblings. Among own siblings, unlike parents' siblings, the younger are marked relative to the older: the three-term "older brother," "older sister," "younger sibling" terminology is much more common the reverse "older sibling," "younger brother," "younger sister" terminology (Murdock, 1970; Nerlove & Romney, 1967). (OT and sibling terminology are the subject of another paper in progress.) This difference in the treatment of parents' siblings and own siblings is not unexpected. A parent's siblings are one generation senior to Ego, whether they are younger or older than the parent, and even a parent's younger siblings are usually older than Ego.

Empirically realized aunt terminologies are only a small subset of possible terminologies. There are 52 possible ways to categorize the five terms M, MoZ, MyZ, FoZ, FyZ, but all but seven possibilities are nonexistent or very rare. In the framework of OT, the work of whittling the possibilities down from 52 to seven is done by just three descriptive constraints (_DLin_, _DBif_, and _DAge_) and two markedness gradients, (_*FZ_ ≥ . _*MZ_ and _*PyZ_ = . _*PoZ_) that limit allowed permutations of classificatory constraints; also, one further markedness gradient, of "aunt" relative to "mother," is implicit in the absence of any _*M_ constraint. These constraints and gradients derive in turn from three schemas of social relationships: schemas of genealogical distance (_DLin_, and the absence of _*M_), of group membership (_DBif_ and _*FZ_ ≥ . _*MZ)_, and of social rank (_DAge_ and _*PyZ_ = . _*PoZ_). The permutations allowed by these constraints and markedness gradients generate the seven major types of aunt terminology without overgenerating possible but nonexistent or very rare types.

3.3. Aunts, uncles, and interdependencies

Not only are empirically realized aunt terminologies a small fraction of the logical possibilities, but many of the same constraints and gradients operate among uncle terminologies as well, generating interdependencies in the categorization of aunts and uncles.

Uncle terminologies follow the same pattern as aunt terminologies, with the sexes interchanged. Go through Table 4 and switch all sexes, replacing “mother” with “father,” “mother’s older brother” with “father’s older sister,” _*FZ_ ≥ . _*MZ_ with _*MB_ ≥ . _*FB_, and so on, and you end up with seven major uncle terminologies, along with associated constraint rankings and markedness gradients that generate them. As in the case of aunt terminologies, all terminologies apart from these seven are nonexistent or very rare (once again ignoring terminologies with sex-of-speaker distinctions). But the variation is even more constrained than this. Aunt and uncle terms do not vary independently.

Table 5 shows how they are related, with aunt terminologies along the top, uncle terminologies along the left, and the number of cases of different combinations given in the corresponding cells. Evident immediately is the concentration of cases along the diagonal. This indicates that aunts and uncles tend to be treated symmetrically: Societies with bifurcate merging terminology for aunts usually have the same for uncles, and so on. In the framework of OT, this implies that *No “parent’s sister”* and *No “parent’s brother”* commonly share a *No*

Table 5
Uncle and aunt interdependencies

	Aunts							
	M	M MZ = FZ	M = MZ FZ	M MZ FZ	M PoZ PyZ	M MoZ MyZ FZ	M MoZ MyZ FoZ FyZ	Other
Uncles								
F	27		1					
F FB = MB	2	73	1[a]	4			1	
F = FB MB	29	3[a]	121	10	[a]	3		3
F FB MB	4	11	9	122	[a]	1		2
F PoB PyB	2	1	[a]	[a]	10	[a]		1
F FoB FyB MB	2	1	9	12	4[a]	56	4	1
F FoB FyB MoB MyB					1	4	9	2
Other			2	3				15

Cells show the number of cultures with different combinations of uncle and aunt terms, from Murdock (1970).

[a] Inconsistent constant rankings (see text).

Kin type abbreviations:

M	Mother
Z	Sister
B	Brother
F	Father
oZ	Older sister
oB	Older brother
P	Parent
yZ	Younger sister
yB	Younger brother

"parent's sibling" stratum, i.e., $(*PZ\ *PB) \equiv *PG$, where G stands for "sibling." However, this association is just a tendency, not a cultural universal or near-universal; exceptions are fairly common (119 cases). When aunts and uncles do not share a stratum, aunts tend to be relatively marked, as shown by comparing the number of cases below (94) and above (25) the diagonal.

There may also be a more subtle interdependency between aunt and uncle terms. Consider the combination of generational aunts (M = MZ = FZ) and bifurcate merging uncles (F = FB≠ MB), which occurs in 29 cases. This combination can be generated with a single ranking embracing both aunts and uncles:

$$*PZ \gg DBif \gg *PB \gg (DLin\ DAge)$$

This combination, although it splits the *No "parent's sibling"* (**PG*) stratum, giving different ranks to *No "parent's sister"* (**PZ*) and *No "parent's brother"* (**PB*), nevertheless gives a single consistent ranking to *DBif*, *DLin*, and *DAge*. Compare this with another possible combination: lineal aunts (M≠ MZ = FZ) and bifurcate merging uncles (F = FB≠ MB). This combination is *inconsistent*, in that no single ranking of the three descriptive constraints could generate it because the aunt terms require $DLin \gg DBif$ while the uncle terms require the opposite, $DBif \gg DLin$. Only separate descriptive constraints for aunts and uncles, for example, *DBifAunt* and *DBifUncle* in place of *DBif*, could generate such an inconsistent combination.

In Table 5, inconsistent combinations are indicated with asterisks. They are relatively infrequent (eight cases). The expected number of inconsistent cases is 27, assuming we calculate expected frequencies of different combinations in the usual fashion, and then change the numbers to reflect the proportions of above-, on-, and below-diagonal cases. Principles of cognitive economy seem to operate here to ensure that aunt and uncle terminologies, if not symmetrical, are normally at least consistent.

4. Cousins

Terminologies for cousins offer more possibilities than those for aunts and uncles because there are more ways to be connected with cousins and more close relations with whom they could be equated. But only a fraction of the possibilities is realized. Actual cousin terminologies fall into six basic types, and many possible types are rare or nonexistent. Cousin terms not only show interdependencies with aunt and uncle terms, but also provide further insight into the cognitive foundations of kin classification.

4.1. From aunts and uncles to cousins

Four of the six basic cousin terminologies correspond to the four basic aunt and uncle terminologies. Eskimo terminologies (74 cases out of 566), including English, which distinguish cousins from siblings, correspond to lineal aunt and uncle terminologies that distinguish lineal (parents) and collateral (parents' siblings) in the parents' generation.

Dravidian/Iroquois terminologies (167 cases) involve a different distinctive feature: they distinguish between parallel cousins, related through parents of the same sex ("mother's sister" and "father's brother"), and cross cousins, related through parents of opposite sex ("mother's brother" and "father's sister"). The former are equated with siblings, while the latter are classified separately as cousins. These terminologies correspond to bifurcate merging aunt and uncle terminologies. (Dravidian and Iroquois differ in how they extend the parallel/cross distinction to second cousins.) In Sudanese terminologies (24 cases), both the lineal/collateral distinction and distinctions based on sex of linking relatives are active: Siblings, parallel cousins, and cross cousins all get separate terms. Sometimes, further distinctions are made between cousins on father's and mother's sides as well. Sudanese terminologies correspond to the maximally descriptive bifurcate collateral aunt and uncle terminologies. Finally, Hawaiian terminologies (171 cases), which classify all cousins and siblings together, correspond to the maximally classificatory generational aunt and uncle terminologies.

These four cases can be handled in the framework of OT by a straightforward extension of the analysis of basic aunt and uncle terminologies. The descriptive constraint _DLin_, which forbids equating parents' siblings with parents, can also be interpreted to rule out equations of cousins with siblings. The descriptive constraint _DBif_, which rules out equations of maternal and paternal parents' siblings can also be interpreted to rule out equations of parallel and cross cousins. Finally, the _No "aunt"_ (i.e., $*\mathit{PZ} \equiv (*\mathit{FZ}\ *\mathit{MZ})$) and _No "uncle"_ (i.e., $*\mathit{PB} \equiv (*\mathit{MB}*\mathit{FB})$) constraint stratum can be supplemented with a *_Cousin_ stratum. (For the present analysis, limited to basic types, a single *_Cousin_ stratum suffices. This stratum could be split up and its components moved around separately to generate the minority of systems that mix different basic types.) With these adjustments, the constraint rankings that generate different aunt and uncle terminologies also generate corresponding cousin terminologies. Just as $\mathit{DLin} \gg *\mathit{PZ} \gg \mathit{DBif}$ generates lineal aunt terminology, so $\mathit{DLin} \gg *\mathit{Cousin} \gg \mathit{DBif}$ generates Eskimo cousin terms, and so on through corresponding terminologies.

This analysis also readily handles interdependencies and markedness relations between aunt/uncle terms and cousin terms. D'Andrade (1971) demonstrates that across cultures cousin terms are commonly found in conjunction with corresponding aunt and uncle terms, but also that where there are discrepancies, cousins are marked relative to parents' siblings. This can be accommodated with a near-universal markedness gradient $*\mathit{Cousin} \geq .\ (*\mathit{PZ}\ *\mathit{PB})$. For example, this gradient permits the ranking:

$$*\mathit{Cousin} \gg \mathit{DBif} \gg *\mathit{PG} \gg \mathit{DLin},$$

which generates the fairly common combination of Hawaiian cousin terms ("cousin" = "sibling") and bifurcate merging aunt and uncle terms (M = MZ ≠ FZ, F = FB ≠ MB). But it rules out the virtually nonexistent converse pattern, in which parents and parents' siblings are merged, while cross cousins are distinguished from siblings. The $*\mathit{Cousin} \geq .\ (*\mathit{PZ}\ *\mathit{PB})$ gradient derives independent support from two schemas of social relationships: the schema of genealogical distance, according to which distant kin are marked relative to close kin, and the schema of social rank, according to which own generation kin are marked relative to ascending generation kin.

4.2. Cousins, generations, and descent lines

Up to this point, the analysis of cousin terms has followed the analysis of aunt and uncle terms. The remaining two basic cousin terminologies, Omaha and Crow, can also be handled by OT, with refinements that shed further light on the psychology of kinship. One of the strengths of OT is that markedness constraints like **Cousin* do not dictate *how* marked forms are avoided when they are avoided. Although the preceding discussion considers only two candidate outputs, "cousin" and "sibling," for the various cousin inputs, OT allows — indeed requires — us to consider candidates from other generations, such as "father," "mother's sister," or "sister's child." Such equations between generations characterize Omaha and Crow cousin terminologies. In Omaha cousin terminologies, "mother's brother's child" is equated with ascending generation relatives on the mother's side of the family, such as "mother," "mother's sister" or "mother's brother." The terminology is normally reciprocal, so that anyone referred to as "mother" by Ego, regardless of the exact genealogical connection, will refer to Ego as "son" or "daughter"; similarly, any classificatory "mother's sister" or "mother's brother" of Ego will refer to Ego as "sister's child." Terminologically speaking, in other words, cousins on the MB side of the family are bumped up a generation, and, reciprocally, cousins on the FZ side are bumped down — equated with descending generation relatives. Crow cousin terminologies are a sex-reversed version of Omaha terminologies, with relatives bumped up on the FZ side and down on the MB side. Murdock's sample of cousin terminologies notes 48 cases of Omaha cousins, and 40 cases of Crow. (Another 32 cases do not fit any of the six basic types: Cousins can only be referred to as "mother's sister's daughter," "uncle's son," etc. These can be accommodated with a slight modification of classificatory constraints, but will not be further considered here. In 10 cases, data are not available.)

It has long been noted that Omaha and Crow cousin terminologies are associated with patrilineal and matrilineal descent, respectively (Fox, 1967; Radcliffe-Brown, 1952; Stone, 2000). Fig. 2 shows how this works in the Omaha case, with four numbered lines of patrilineal descent associated with Ego and Ego's siblings, parents, parents' siblings, and cousins. (One might think of the Roman numerals as corresponding to surnames, because surnames commonly descend patrilineally in US American culture.) By equating MBC with M, MB, or MZ, Omaha terminologies emphasize that MBC, unlike any other cousin, is part of mother's natal patriline, line III. Roughly speaking, terms for "mother," "mother's sister," and "mother's brother" take on an extended meaning something like "female or male of mother's patriline." Note that the discussion is framed in terms of descent *lines*, not descent *groups*. The two are related — societies that emphasize patrilineal lines of descent in terminology also tend to have socially important patrilineal descent groups — but the association between terminology and institutions is not perfect (Needham, 1971; Parkin, 1997, pp. 72–76).

The case of matrilineal descent, associated with Crow terminologies, is shown in Fig. 3. Here, there are four numbered lines of matrilineal descent, and the equation of FZC with F, FZ, or FB emphasizes that FZC is part of father's matriline, line II.

To accommodate these patrilineal and matrilineal distinctions within the framework of OT requires several elaborations of the analysis of cousin terms given above. Because some cousin

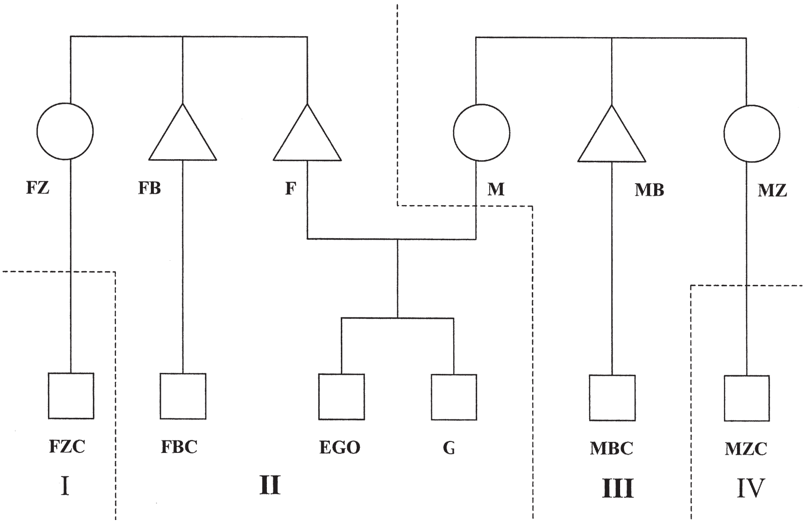

Fig. 2. Adjacent patrilineages.

terminologies equate cousins with relatives in other generations, while others do not, it is necessary to introduce a descriptive constraint, *DGen*, that keeps different generations separate unless overruled by higher ranking constraints. And because Omaha and Crow terminologies do not merely distinguish between cross and parallel kin, but also exhibit patrilineal and matrilineal skewing, it is necessary to consider both patrilineal and matrilineal versions of the bifurcatory constraint *DBif*. Below, I introduce and compare generational and unilineal

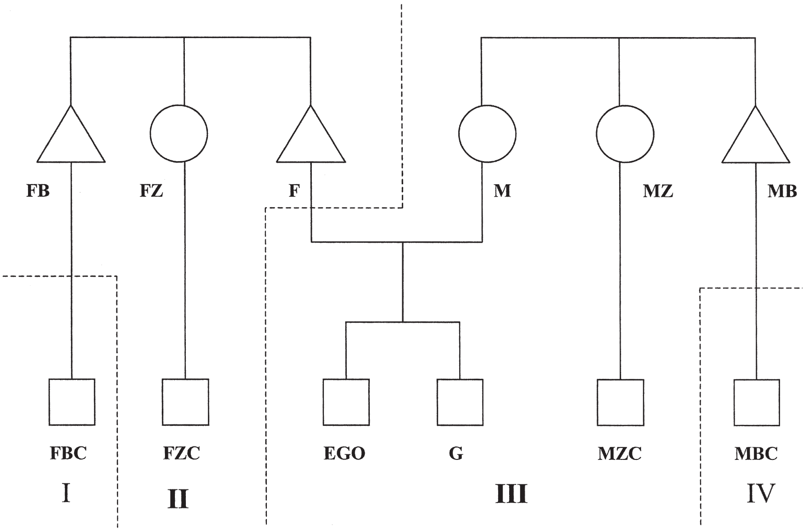

Fig. 3. Adjacent matrilineages.

constraints and then show how the revised set of constraints generates not just Omaha and Crow terminologies but also the four basic cousin terminologies considered above.

4.2.1. DGen

All kin terminologies observe some distinctions between relatives in different generations, implying one or more descriptive constraints related to generations. In an analysis of terms for grandparents and grandchildren (not presented here), I find evidence for two such constraints. One requires that adjacent generations be kept distinct — own generation from parental generation, parental from grandparental, and so on — but allows mergers involving nonadjacent (reciprocal or alternating) generations. The other constraint requires that relatives in ascending and descending generations be kept distinct: If the input is a relative in the ascending generation, the output must be as well. Varying the rank order of these constraints along with markedness constraints generates various common patterns, including "grandparent" = "parent" and "grandchild" = "child" equations, and interdependencies and equations between grandchildren and grandparents. For present purposes, in which only the $+1$, 0, and -1 generations are considered, these two constraints are equivalent and can be reduced to a single constraint, *DGen*, which simply requires that input and output be of the same generation. However, a more complete analysis of Omaha and Crow terminologies, including candidate outputs in the $+2$ and -2 generations, i.e., grandparents and grandchildren, would require additional machinery.

4.2.2. DPatri and DMatri

In Omaha cousin terminologies, MBC, belonging to mother's patriline (III), is equated with ascending generation members of patriline III, and distinguished from cousins in other patrilines. This suggests a descriptive constraint, here labeled *DPatri*, which enforces distinctions between patrilineal descent lines. But *DPatri* must do something more complicated than just demand that input and output belong to the same patriline, because Omaha terminologies commonly *do* equate some cousins from different patrilines, like the two parallel cousins, FBC and MZC (patrilines II and IV), who are both commonly equated with siblings. We can capture the principle that seems to be involved here by introducing the concept of *adjacent* or *cross* patrilines: Two patrilines are *adjacent* if they are the patrilines of somebody's mother and father. Adjacent patrilines are spatially adjacent in Fig. 2. Then *DPatri* can be defined as forbidding Input → Output transformations between adjacent patrilines, but not between more distant ones. If MBC (line III) is the input, then *DPatri* prohibits any relative in II as output, because MBC belongs to Ego's maternal patriline (the patriline of Ego's mother), while II is Ego's paternal patriline (the patriline of Ego's father — and of Ego). Similarly, if FZC (line I) is the input, then the output cannot be any relative from Ego's patriline (line II), because this would violate the separation between paternal and maternal patrilines of FZC. On the other hand, inputting MZC (line IV) and getting "sibling" (line II) as output does not violate *DPatri*, because Ego's patriline is not adjacent to MZC's: There is nobody for whom IV and II are maternal and paternal patrilines. Thus, the forbidden pairs of adjacent patrilines are I and II, II and III, and III and IV, but not I and III, I and IV, or II and IV.

A matrilineal version of the separation of adjacent descent lines seems to be at work in Crow terminologies. Here, another descriptive constraint, *DMatri*, separating adjacent matrilineal rather than patrilineal lines of descent, is important.

Generational and unilineal distinctions are worth comparing briefly. First, in both cases, there are constraints that separate adjacent or cross categories (adjacent generations or lines) while allowing equations between more distant categories. There may be a parallel with phonology, where the principle that adjacent syllables should be differentially stressed can result in patterns of alternating stress (Jackendoff, 2002). All of these cases involve something like the cartographer's rule that adjacent countries on a map should be colored differently to emphasize their distinctiveness. This rule, together with a limited stock of colors and the facts of geography, may require that two nonadjacent countries with nothing special in common be colored the same (Aberle, 1967; Hage, 1999).

Second, in the case of generations, another principle is at work, one that partitions ascending from descending generations. If we sought a parallel in the case of collateral distinctions, it would be a descriptive constraint based on laterality rather than unilineality, a constraint partitioning relatives on the father's side, parallel or cross, from relatives in the mother's side. If this patrilateral/matrilateral distinction were important, we would expect to find many terminologies making no parallel/cross distinction, but having one term for patrilateral cousins, FBC and FZC, and another for matrilateral cousins, MZC and MBC. Yet, such systems are rare or nonexistent. The laterality distinction seems to be important not so much in its own right, but in making finer distinctions among patrilines or among matrilines. In other words, the distinctions between ascending and descending generations and among patrilines and matrilines are more *salient* than the distinction between patrilaterality and matrilaterality. (Greenberg, 1990 observes that just as some kin types are more marked than others, so some dimensions of kin classification are more salient than others.) The low salience of the laterality distinction is predictable on the supposition that constraints and constraint gradients derive, not from the logical possibilities for categorizing kin, but from schemas of social relationships. The ascending/descending distinction derives its salience from the schema of social rank, the unilineal distinctions from the schema of group membership.

Table 6 gives constraints and constraint rankings associated with different basic cousin terminologies. Table 7 shows these constraints in action, generating an Omaha cousin terminology, with three tableaus for different cousin type inputs based on the ranking:

$$(^{*}\mathit{Cousin}\ \mathit{DPatri}) \gg (\mathit{DGen}\ \mathit{DMatri}) \gg \mathit{DLin}.$$

In the first tableau, FBC is the input. The high rank of the **Cousin* constraint rules out any type of cousin as output. (As noted in a footnote in Table 7, I combine all cousin candidates into one row, because different types' violations of lower ranking constraints do not affect the outcome.) Among the other candidates, only G (sibling) satisfies the next two constraints, because FBC and G both belong to Ego's patriline and generation. In the second tableau, MZC too gives "sibling" as the optimal output, because this transformation, although it entails equating relatives from different patrilines, nonetheless respects the distinction

Table 6
Generating cousin terms

Type	Equations	Constraint ranking
Omaha	Patrilineal skew	(_*Cousin DPatri_) ≫ (_DGen DMatri_) _DLin_
Crow	Matrilineal skew	(_*Cousin DMatri_) ≫ (_DGen DPatri_) _DLin_
Dravidian/Iroquis	Sibling = ‖cousin Xcousin	(_DPatri DMatri_) ≫ (_*Cousin DGen_) _DLin_
Sudanese	Sibling ‖cousin Xcousin	(_DPatri DMatri DGen_) ≫ (_DLin_)_* Cousin_
Eskimo	Sibling ‖cousin Xcousin	_DGen DLin*Cousin_ ≫ (_DPatri DMatri_)
Hawaiian	Sibling = ‖cousin = X cousin	(_DGen *Cousin_) ≫ (_DLin DPatri DMatri_)

Kin type abbreviations:

‖cousin	Parallel cousin	Xcousin	Cross cousin

Descriptive constraints:

DPatri	_Distinguish adjacent patrilines_
DMatri	_Distinguish adjacent matrilines_
DGen	_Distinguish generations_
DLin	_Distinguish lineal and collateral kin_
DLat	_Distinguish matrilateral and patrilateral kin_ Note: The matrilateral/patrilateral distinction does not distinguish major types listed above but sometimes plays a secondary role in subtypes that draw further distinctions among patrilineages or among matrilineages.

Classificatory constraint:

*Cousin ≡ (_*FZC *MBC *FBC *MZC_)	_No "cousin"_ stratum equivalent to (_No "father's sister's child" No "mother's brother's child,"_ etc.) Note: The _No "cousin"_ stratum may be split up in mixed cousin terminologies, or terminologies that apply sex or other distinctions to some but not all cousins.

Gradients:

*Cousin ≥ . (_*PZ *PB_)	Cousin marked relative to aunt and uncle
(_DPatri DMatri_) ≥ . _DLat_	Unilinearity more salient than laterality
DGen ≥ . _DLin_	Generation more salient than lineal/collateral distinction

between own/paternal and maternal patrilines. But MBC, in the third tableau, cannot be equated with "sibling," because this violates the adjacent patriline constraint. Some alternative must be found. As long as the generation distinction, *DGen*, is ranked low, optimal outputs can include ascending generation relatives in Ego's maternal patriline, such as MB, MZ, or M, with the exact output depending on sex of relative and the ranking of *DLin* relative to _*Aunt_ (omitted here).

This leaves one final cousin type, FZC. Omaha terminologies normally handle FZC by making use of the fact that "father's sister's child" and "mother's brother's child" are reciprocals: Ego is FZC to her MBC, and vice versa. Therefore, if Ego refers to her MBC as "mother's brother" or "mother," then this person will refer to Ego, her or his FZC, by the reciprocal terms, "child" or "sister's child" (Fox 1967, pp. 252–256; Lounsbury, 1964). Working with reciprocal terms requires a revision of OT, allowing descriptive constraints to operate not only between Input and Output but also between, for example, Reciprocal of Input and Reciprocal of Output. The necessary modifications are introduced in Appendix.

The analysis above represents a formalization of the common informal observation that Omaha terminological equations result from letting patrilineal principles take precedence over

Table 7
Omaha cousins: tableaus
Father’s brother’s child (FBC) is “sibling” (G) . . .

FBC	**Cousin*	*DPatri*	*DGen*	*DMatri*	*DLin*
PGC (cousin)	*!	[a]		[a]	
MB		*!	*		
FB			*!	*	
G✓					*
M		*!	*		*
F			*!	*	*

and mother’s sister’s child (MBC) is “sibling” (G) . . .

MZC	**Cousin*	*DPatri*	*DGen*	*DMatri*	*DLin*
PGC (cousin)	*!	[a]		[a]	
MB		*!	*		
FB			*!	*	
G✓					*
M		*!	*		*
F			*!	*	*

but mother’s brother’s child (MBC) is “mother’s brother” (MB) or “mother” (M) (depending on his or her sex, not shown here) . . .

MBC	**Cousin*	*DPatri*	*DGen*	*DMatri*	*DLin*
PGC (cousin)	*!	[a]		[a]	
MB✓			*	*	
FB		*!	*		
G		*!		*	*
M✓			*	*	*
F		*!	*		*

so parallel cousins (FBC, MZC) are equated with siblings, while MBC is equated with ascending generation kin in the patrilineage. For the fourth cousin type, FZC, see discussion of reciprocal terms in Appendix.

See Table 6 for definitions and abbreviations.

[a] These constraints may or may not be violated, depending on the specific cousin type, but this doesn’t affect the optimal output.

generational and other principles. Other constraint rankings generate the other basic cousin types, as set forth in Table 6, and discussed below.

Crow terminologies are a sex-reversed version of Omaha terminologies, with _DPatri_ and _DMatri_ interchanged:

$$(*\mathit{Cousin}\ \mathit{DMatri}) \gg (\mathit{DGen}\ \mathit{DPatri}) \gg \mathit{DLin}.$$

The remaining terminologies, which treat patrilines and matrilines symmetrically or not at all, are generated by assigning the two unilineal distinctions to a common stratum, _DBif_ ≡ (_DPatri DMatri_), and permuting the rankings of _DBif_, _DLin_ and _*Cousin_. For example, Table 8 shows two tableaus for Dravidian/Iroquois terminology, corresponding to the ranking:

$$(\mathit{DPatri}\ \mathit{DMatri}) \gg (*\mathit{Cousin}\ \mathit{DGen}) \gg \mathit{DLin}.$$

Applied to cross cousin inputs (first tableau), the high rank of the _DBif_ ≡ (_DPatri DMatri_) stratum prevents cross cousins from being equated with other relatives, while _*Cousin_ ensures that they merge with one another. (The actual output, FZC or MBC, depends on the order of constraints within the _*Cousin_ stratum, and is not shown here). Applied to parallel cousin inputs (second tableau), the ranking (_DPatri DMatri_) $\gg$ _*Cousin_ produces equations with siblings.

Table 6 also includes three universal or near universal gradients. The first two were discussed previously. One is a markedness gradient: cousins are marked relative to aunts and uncles, _*Cousin_ $\geq .$ (_*PZ *PB_). The second is a gradient in salience: Unilineality is more salient than laterality, (_DPatri DMatri_) $\geq .$ _DLat_. The laterality distinction _DLat_ is mostly omitted from the constraint rankings and tableaus below, but acts as a tiebreaker in the discussion of reciprocal terms in Appendix and Table 9. The third gradient, _DGen_ $\geq .$ _DLin_, states that generational distinctions are more salient than lineal/collateral ones. Consider the fairly common highly classificatory terminology generated by the ranking:

$$*\mathit{Cousin} \gg (\mathit{DGen}\ *\mathit{PG}) \gg (\mathit{DLin}\ \mathit{DPatri}\ \mathit{DMatri}).$$

In this terminology, there are no separate terms for cousins or parents' siblings; the former are equated with siblings, the latter with parents. But if, in violation of _DGen_ $\geq .$ _DLin_, _DLin_ and _DGen_ were interchanged, then all collaterals — cousins, parents' siblings, and probably siblings' children — would be equated with one another and distinguished from lineal relatives. Such extreme collateral merging, in contrast with extreme generational merging, is rare or nonexistent. The greater salience of generational distinctions is consistent with the fact that these involve two schemas of social relationships, social rank and genealogical distance, while the lineal/collateral distinction involves only the latter.

The analysis above is satisfying because it uses a few principles to capture important regularities, generating major types without overgenerating. It also has some implications about the sociological and cognitive bases of kin classification, especially the parallel/cross

Table 8
Dravidian/Iroqois cousins: tableaus
Mother's sister's child (MZC) is "sibling" (G) . . .

MZC	*DPatri*	*DMatri*	**Cousin*	*DGen*	*DLin*
PGC (cousin)	[a]	[a]	*!		
MB	*!			*	
FB		*!		*	
G✓					*
M	*!			*	*
F		*!		*	*

and mother's brother's child (MBC) is "cousin" (MBC or FZC, depending on order within the **Cousin* stratum, not shown here) . . .

MBC	*DPatri*	*DMatri*	**Cousin*	*DGen*	*DLin*
MBC or FZC✓			*		
MZC	*!	*	*		
FBC	*!		*		
MB		*!		*	
FB	*!			*	
G	*!	*			*
M		*!		*	*
F	*!			*	*

and, generally, parallel cousins (MZC, FBC) are equated with siblings but distinguished from cross cousins (MBC, FZC).

See Table 6 for definitions and abbreviations.

[a] These constraints may or may not be violated, depending on the specific cousin type, but this does not affect the optimal output.

distinction. In some cultures, this distinction is a straightforward reflection of social facts. This is the case especially where Dravidian terminologies are found together with a rule of cross cousin marriage or moiety exogamy. In these societies, parallel cousins, classified as siblings, are prohibited as marriage partners, while cross cousins are allowed or preferred spouses. These rules may be further emphasized by terminological equations between cross relatives and affines, with Ego's opposite sex cross cousins called "husband" or "wife," cross aunts and uncles (FZ and MB) called "spouse's mother/father," and so on (Dumont, 1983).

Yet, parallel/cross distinctions are also common in societies that merely emphasize patrilineal or matrilineal descent but lack the corresponding systems of marriage alliance.

In these cases, terminology is not just a map of social structure: The two types of parallel cousin, MZC and FBC, may both be equated with siblings even though they are not all members of the same lineage. What the present analysis suggests is that these patterns may arise as a byproduct of principles of cognitive economy. As long as there is a tendency to treat the two sexes symmetrically through a *DBif* ≡ (*DPatri DMatri*) stratum, then emphasis on either patrilineality or matrilineality, with a correspondingly high rank for *DBif*, can result in parallel/cross distinctions in terminology. There is a parallel here with the tendency, shown in Table 5, to treat aunt and uncle terminology symmetrically, reflecting a widespread **PG* ≡ (**PZ *PB*) stratum. On this account, kin terminology reflects a combination of sociological influences, which affect the rank of different constraints, and cognitive principles, which determine both how constraint rankings generate terminologies and what constraints and strata are available in the first place.

5. Taking stock

Why have many linguists taken strongly to OT? OT subsumes earlier work on rules of language in a more inclusive framework. The rules noted by earlier researchers are still present, but as surface phenomena — products of particular constraint rankings — while a universal repertoire of constraints, gradients, and schemas explains why some rules and not others are allowed. One result is a satisfying account of *conspiracies* among rules. Consider a familiar phenomenon from phonology: Evidence from many languages shows that the prototypical, maximally unmarked syllable is a so-called "CV" (consonant–vowel) syllable, consisting of a single consonant onset followed by one vowel. *Sigh* and *key* are CV syllables, while *egg* is VC, *string* is CCCVC, and *sixths* is CVCCCC. The preference for syllables with single-consonant onsets expresses itself in many ways. In English, some multiconsonant syllable onsets are avoided while others are tolerated. Some languages avoid all such onsets, others tolerate them more freely than English. And languages follow different strategies in avoiding multiconsonant onsets: Sometimes, consonants are deleted, other times, vowels are added. It is possible to come up with a collection of formal substitution rules that cover the various possibilities in different languages: rules for deleting all surplus consonants, rules for deleting them only in specific phonological environments, rules for adding vowels sometimes, or always. But what is missing in such a collection of rules is a formal account of what is evident informally: All the different rules aim or conspire toward a single end, the avoidance of multiconsonant syllable onsets. In OT, by contrast, this avoidance is part of the formal machinery itself. Multiconsonant onsets are banned by a single markedness constraint, while the surface effects of this constraint vary depending on its ranking and on how its expression is modified or overridden by other constraints. Constraints in OT are more flexible than substitution rules because there is nothing in a constraint to specify how or to what extent it is satisfied. It is only the interaction of one constraint with others, filtering variation rather than applying set transformations, that dictates the output.

OT offers the same advantage when applied to kin terminology. Previous formal approaches to kin terminology, like earlier linguistic work, have often relied on substitution rules to account for mergers of kin terms. For example, a rule allowing the replacement of "mother's sister" with "mother" generates a bifurcate merging aunt terminology ($M = MZ \neq FZ$). Other substitution rules can generate corresponding uncle terminologies; still others generate the mergers of sibling and cousin terms characteristic of Hawaiian, Iroquois, and Dravidian terminologies, and the mergers across generations characteristic of Omaha and Crow terminologies. (Woolford, 1984 provides an important review and synthesis.) Absent from such an account, however, is what is apparent informally: A whole set of rules conspires toward the same end, recognizing distinctions among unilineal descent lines. In OT, by contrast, two constraints directly related to the separation of patrilines and of matrilines, *DPatri* and *DMatri*, generate a wide range of surface effects depending on their rank and interaction with other constraints.

Not only does OT have advantages as a formal approach to kin classification, but it also provides a new perspective on old debates, including debates about sociological vs. cognitive approaches to kin classification, and about domain-specific vs. domain-general theories of cognition. Consider the first debate: Ever since Morgan, anthropologists have proposed that different kin terminologies correspond to different social institutions, and these proposals have often been supported by cross-cultural correlational studies (Murdock, 1949; Stone, 2000; Whiting, 1988). Yet, the sociological approach to kin terminology has been countered by other anthropologists and linguists (Greenberg, 1966, 1990; Kroeber, 1909; Lounsbury, 1964) who have noted that not only are these correlations imperfect, but also that many recurring patterns in kin classification do not plausibly map onto social institutions, and seem better explained by cognitive factors. The OT approach to kin classification suggests a synthesis of the two positions. Institutions and social roles do not translate directly into terminology. Instead, the importance of different principles of social organization, like patrilineality and matrilineality, translates roughly into an underlying relative ranking of different constraints, like *DPatri* and *DMatri*, and this ranking in turn generates a particular terminology, with markedness effects and so on, according to the principles of OT.

Furthermore, OT combined with evolutionary psychology suggests that at the level of universal psychology sociological and cognitive explanations are not as far apart as they otherwise seem. In any one society, social institutions and categories may not line up perfectly. But across the species, the same underlying principles show up repeatedly in both sociological and linguistic domains, because both are products of the same generative psychology.

Finally, OT offers a new perspective on debates over domain-specific vs. domain-general cognition. On the one hand, many cognitive scientists and evolutionary psychologists argue that human cognition is domain-specific: that the human mind comprises a collection of psychological mechanisms or mental organs specialized for different domains of learning and reasoning. OT is consistent with this approach, and in the companion paper, I argue that OT applied to kinship provides evidence for three evolved schemas of social relationships.

On the other hand, there is another tradition in the study of language and mind, a tradition of looking for general principles of thought that produce similarities rather than differences across cognitive domains. This tradition is represented by structural linguists like Greenberg (1975) and, in a more speculative vein, by structuralist anthropologists like Lévi-Strauss

(1967, 1969). Recurring phenomena like binary distinctive features, linguistic markedness, prototype effects, and implicational hierarchies hint that general principles are at work, at least in the more systematic and rule-governed areas of cognition (Pinker, 1999; Pinker and Prince, 1996). If OT can prove itself not only in phonology and syntax, but also in semantic domains like kinship, this argues that along with domain-specific mechanisms there are also — at right angles as it were — domain-general principles of cognitive or communicational economy that generate regular cultural variation out of trade-offs between binary distinctions and markedness hierarchies.

Just how widely OT can be applied remains to be seen. The analysis in this and the companion paper suggests that when it comes to categorizing kin the theory offers both a formally satisfying approach, and a window on the intersection of evolved social and cognitive psychologies.

Appendix

Kinship terminologies sometimes involve more complex correspondences than the direct input–output relationships discussed so far. Some correspondences involve parts of words. For example, although English has a separate word for "granddaughter" (generation −2), the word is morphologically complex, consisting of a root identical to the adjacent generation term "daughter" (generation −1), and a prefix identical to the prefix for the reciprocal generation term "grandparent" (generation +2). Other correspondences involve parts of long kinship formulas for distant relations, rather than parts of words. For example, constraints may or may not evaluate "cousin" and its potential replacements the same way inside a larger formula as on their own. This is the basis for the distinction between Dravidian and Iroquois cousin terminologies, which treat first cousins the same ("sibling" = "parallel cousin" ≠ "cross cousin"), but "parent's cousin's child" (second cousin) differently. In Iroquois terminologies, any embedded cousins are equated with siblings, so all of mother's male cousins are equated with "mother's brother," and so on; Dravidian terminologies treat parents' cousins differently depending on whether they are parallel or cross.

Another sort of correspondence involves reciprocal rather than direct correspondences between input and output. By definition, the reciprocal of the relationship of A to B is the relationship of B to A, and vice versa. For example, "mother's brother" and "man's sister's child" are reciprocals, as are "mother" and "woman's child." (As these examples suggest, sex-of-speaker distinctions are often a sign of reciprocal correspondences.) In one version of reciprocal dependence, the same term is used for a kin type and its reciprocal. More commonly, two types are equated because the reciprocal of one is equated with the reciprocal of the other. For example, in Omaha cousin terminologies, the general rule is that anyone who calls Ego "mother's brother" (or "mother") should be called "man's sister's child" (or "mother's child") by Ego. Because one cousin type, "mother's brother's child," is equated with "mother's brother" (or "mother"), the reciprocal cousin type, "father's sister's child," is generally equated with the reciprocal of "mother's brother" (or of "mother").

Table 9
Omaha cousins, reciprocal terms: tableau

FZC or MBC^{r}	**Cousin*	*DPatri*	*DGen*	*DMatri*	*DLin*	*DLat*
PGC (cousin)	*!	[a]		[a]		
MB			*	*		*!
MB^{r}✓			*	*		
FB		*!	*			
FB^{r}		*!	*			*
G		*!		*	*	
M			*	*	*	*!
M^{r}✓			*	*	*	
F		*!	*		*	
F^{r}		*!	*		*	*

Father's sister's child (FZC), the reciprocal of mother's brother's child (MBC^{r}), is either "man's sister's child" (mZC), the reciprocal of mother's brother (MB^{r}), or "woman's child" (wC), the reciprocal of mother (M^{r}), depending on his or her sex (not shown here).

[a] These constraints may or may not be violated, depending on the specific cousin type, but this does not affect the optimal output.

All of these complex input–output correspondences can be accommodated in an OT framework by expanding the range of allowed candidate outputs and introducing a unified set of principles for handling them. Here, I briefly indicate how this works for the reciprocal dependences in Omaha cousin terms, illustrated with the tableau in Table 9. Note first that the input is presented in two forms, FZC *or* MBC^{r}.This notation implies that although "father's sister's child" (FZC) and "reciprocal of mother's brother's child" (MBC^{r}) refer to the same kin type, they may be psychologically distinct, with the preferred form dictated by the candidate output. When a candidate output carries no superscript the direct interpretation of the input, FZC, is preferred. For example, the candidate MB activates the Input → Output transformation FZC → MB. Because FZC and MB belong to different generations, to adjacent matrilineages, and to father's and mother's sides of the family, the tableau shows candidate MB violating *DGen*, *DMatri*, and *DLat*, but not other Input–Output constraints.

Another candidate output, MB^{r}, activates reciprocal correspondences. MB^{r} means "man's sister's child" conceived of as "reciprocal of mother's brother." The superscript implies that the transformation to be analyzed is MBC^{r} → MB^{r}, which is treated *as if* it were the transformation MBC → MB. This convention captures the fact that two terms may be equated owing not to their direct similarities but to similarities between their reciprocals. Because MBC and MB belong to different generations and to adjacent matrilineages, MBC^{r} → MB^{r} violates *DGen* and *DMatri*. But MBC and MB are both matrilateral relatives, so MBC^{r} → MB^{r} does not violate *DLat*. Here, *DLat* acts as a tiebreaker, making MB^{r} rather than MB an optimal output: other things being equal, it is better to equate relatives in the same patriline than relatives in nonadjacent but different patrilines. Thus, "reciprocal of mother's brother's child" is equated with "reciprocal of mother's brother."

The argument up to this point applies only when Ego is male. For female Ego, the tableau shows that the optimal output is M^r — "reciprocal of mother," i.e., "woman's child." An analogous argument for Crow cousin terminology leads to "mother's brother's child" being equated with "woman's brother's child" or "man's child." A more complete treatment of reciprocal and other complex input–output correspondences will be presented in a discussion of grandparent–grandchild terminology (Jones, in preparation).

References

Aberle, D. F. (1967). A scale of alternate generation terminology. *Southwestern Journal of Anthropology*, *23*, 261–276.

Archangeli, D. (1997). Optimality theory: an introduction to linguistics in the 1990s. In: D. Archangeli, & D. T. Langendoen (Eds.), *Optimality theory: an overview* (pp. 1–32). Oxford: Blackwell.

Boersma, P., Dekkers, J., & van de Weijer, J. (2000). Introduction. In: J. Dekkers, F. van der Leeuw, & J. van de Weijer (Eds.), *Optimality theory: phonology, syntax, and acquisition* (pp. 1–44). Oxford: Oxford University Press.

Cziko, G. (1995). *Without miracles: universal selection theory and the second Darwinian revolution*. Cambridge: MIT Press.

D'Andrade, R. G. (1971). Procedures for predicting kinship terminology from features of social organization. In: P. Kay (Ed.), *Explorations in mathematical anthropology* (pp. 60–75). Cambridge: MIT Press.

Dennett, D. (1995). *Darwin's dangerous idea: evolution and the meanings of life*. New York: Simon and Schuster.

Dumont, L. (1983). *Affinity as a value: marriage alliance in South India, with comparative essays on Australia*. Chicago: University of Chicago Press.

Fox, R. (1967). *Kinship and marriage: an anthropological perspective*. Harmondsworth, Middlesex: Penguin.

Goodenough, W. (1970). *Description and comparison in cultural anthropology.* Chicago: Aldine.

Greenberg, J. (1975). Research on language universals. *Annual Review of Anthropology*, *4*, 75–94.

Greenberg, J. (1990). Universals of kinship terminology: their nature and the problem of their explanation. In: K. Denning, & S. Kemmer (Eds.), *On language: selected writings of Joseph Greenberg*. Stanford: Stanford University Press.

Greenberg, J. H. (1966). *Language universals, with special reference to feature hierarchies*. The Hague: Mouton.

Hage, P. (1997). Unthinkable categories and the fundamental laws of kinship. *American Ethnologist*, *24*, 652–667.

Hage, P. (1999). Alternate generation terminology: a theory for a finding. *Journal of Anthropological Research*, *55*, 521–539.

Jackendoff, R. (2002). *Foundations of language: brain, meaning, grammar, evaluation*. Oxford: Oxford University Press.

Jones, D. (in preparation). Optimality theory and kin classification: siblings and grandkin.

Kroeber, A. L. (1909). Classificatory systems of relationship. *Journal of the Royal Anthropological Institute*, *39*, 77–84.

Levi-Strauss, C. (1967). *Structural anthropology.* Garden City, NY: Anchor Books, Doubleday.

Levi-Strauss, C. (1969). *The elementary structures of kinship*. Boston: Beacon Press.

Lounsbury, F. (1964). A formal account of the Crow- and Omaha-type kinship terminologies. In: W. H. Goodenough (Ed.), *Explorations in cultural anthropology: essays in honor of George Peter Murdock* (pp. 351–393). New York: McGraw-Hill.

Morgan, L. H. (1870 [1997]). *Systems of consanguinity and affinity of the human family.* Lincoln: University of Nebraska Press.

Murdock, G. P. (1949). *Social structure*. New York: MacMillan.

Murdock, G. P. (1970). Kin term patterns and their distribution. *Ethnology*, *9*, 165–207.

Nerlove, S., & Romney, A. K. (1967). Sibling terminology and cross-sex behavior. *American Anthropologist*, *69*, 179–187.

Parkin, R. (1997). *Kinship: an introduction to basic concepts*. Oxford: Blackwell.

Pinker, S. (1999). *Words and rules: the ingredients of language*. New York: Basic Books.

Pinker, S., & Prince, A. (1996). The nature of human concepts: evidence from an unusual source. *Communication and Cognition*, *29*, 307–361.

Prince, A., & Smolensky, P. (1997). *Optimality Theory: constraint interaction in generative grammar*. Report no. RuCCS-TR-2. New Brunswick, NJ: Rutgers University Center for Cognitive Science.

Prince, A., & Smolensky, P. (1997). Optimality: from neural networks to universal grammar. *Science*, *275*, 1604–1610.

Radcliffe-Brown, A. R. (1952). *Structure and function in primitive society: essays and addresses*. Glencoe, IL: The Free Press.

Stone, L. (2000). *Kinship and gender: an introduction*. Boulder, CO: Westview Press.

Tesar, B., Grimshaw, J., & Prince, A. (1999). Linguistic and cognitive explanation in Optimality Theory. In: E. Lepore, & Z. Pylyshyn (Eds.), *What is cognitive science?* (pp. 295–326). Malden, MA: Blackwell.

Woolford, E. (1984). Universals and role options in kinship terminology: a synthesis of three formal approaches. *American Ethnologist*, *11*, 771–790.