Topics in Linguistics (2023), 24(2), pp. 15-29

10.2478/topling-2023-0009

The semantic complexity of Hausa kinship terms

Gian Claudio Batic*
University of Naples L'Orientale, Italy

Abstract
This study aims at analysing the absolute semantic complexity of kin terms in Hausa, i.e. to measure the amount of semantic information of individual kin terms. Each kin term is defined by a set of sufficient and necessary conditions (i.e. properties and relations) derived from the construction of a genealogical "space". In order to calculate semantic complexity, properties (e.g. *x is male*, *x is older than y*) and relations (e.g. *x is married to y*, *x is father of y*) are encoded as a series of predicates. The terms are defined in a feature matrix system: for each property and relation each kin term is assigned a value on a truth table. Resorting to predicate calculus, the complexity coefficient *c* of kin terms is calculated as the negative dyadic logarithm of the relative number of trues according to the formula proposed by Lehmann (1978) and adapted from Carnap and Bar-Hillel (1952). Being culture-independent, the definition of kinship terms in a feature-matrix system allows for a) cross-linguistic comparison; b) a consistent treatment of polysemous instances based on the principles of intension and extension; and c) further analysis and applications in representations of kinship systems formulated with genealogical or algebraic approaches.

Key words
Hausa, kinship terminology, complexity, predicate calculus, polysemy

1. Introduction
This article aims to address the following questions: a) What is the amount of semantic information contained in Hausa kinship terms? b) How is semantic complexity distributed in Hausa kinship terminology? The two questions are embedded in the more general debate on "local" linguistic complexity (i.e. pertaining to certain areas of grammar and lexicon) and are relevant for the purpose of cross-linguistic comparison. The article is structured as follows: Section 2 illustrates the terminology of Hausa kinship, with a comparison of some culture-specific terms as described in existing lexical compilations. Section 3 is devoted to the definition of linguistic complexity and the methodological principles adopted, while Section 4 outlines the principle underlying the treatment of polysemous instances. Section 5 illustrates the formal coding of features (properties and relations) that define kin terms, and Section 6 offers a detailed description of the calculation procedure used to determine semantic complexity values. Section 7 presents the complexity values of each term, followed by a discussion of the results in Section 8.

2. Hausa kinship terms
Research into kinship structures was one of the earliest interests of anthropological science. Morgan's pioneering work, *Systems of Consanguinity and Affinity of the Human Family* (1870), proposed a first

* Address for correspondence: Gian Claudio Batic, Department of Asian, African and Mediterranean Studies, University of Naples L'Orientale, Piazza San Domenico Maggiore 12, Naples, Italy. E-mail: gcbatic@unior.it

classification of kinship systems, outlining a typology based on classificatory and descriptive systems. Morgan's typology was developed and refined over the next 80 years by Spier (1925), Lowie (1928), Kirchhoff (1932), and Murdock (1949), and the result was a typology comprising six system types, i.e. the well-known Hawaiian, Sudanese, Omaha, Eskimo, Crow, and Iroquois types. Although the study of kinship structures has progressed by accumulating data and refining the theoretical framework, most of the kinship systems of the human family (with a few exceptions) can be ascribed to one of Morgan-Lowie-Murdock's six system types. As Bernard points out, "anthropologists also noticed very early that, although kinship systems could be unique to each culture [...] they simply weren't" (Bernard 2011: 223-224).[2]

If we consider the criteria of the classical typology, Hausa kinship terminology falls only partially within the Sudanese system, a system that assigns a term to almost each member of Self's[3] kin based on distance from Self, gender and type of relationship. A hallmark feature of the Sudanese kinship system is the distinction between parallel cousins, maternal cross-cousins, and paternal cross-cousins. The Hausa system, on the other hand, is descriptive in the parental generation (i.e. it differentiates between the mother's siblings and the father's siblings), while in the Self generation it is characterized by the distinction between parallel cousins and cross-cousins and by a certain degree of overlap between siblings and cousins. In other words, as we shall see below, Hausa kinship terminology employs both classificatory and descriptive terms.

The notion of kinship among the Hausas is encoded by two complementary terms: *dangi* and *iyali*. *Dangi* indicates the network of collateral and in-law relationships in which Self is embedded. The term includes also the lineal intergenerational bonds, but excludes the terms referring to parents and children. *Iyali*, on the other hand, is the term reserved for the parents-children subset. In the present analysis, the whole relational complex *dangi-iyali* will be considered.

The Hausa kinship terms, subdivided by generation and linearity, are indicated below. Non-epicene and non-gender-specific terms are given in pairs *masculine* | *feminine*.

Ascending
(i) *uba* 'father' | *uwa* 'mother'
(ii) *mahaifi* 'male parent' | *mahaifiya* 'female parent'
(iii) *kaka* 'grandfather, grandmother'

Ascending lateral
(i) *baba/baffa* 'father's brother (younger or elder)' | *goggo/gwaggo* 'father's sister (younger or elder)'
(ii) *rafani* 'mother's brother' | *inna* 'mother's younger or elder sister'
(iii) *kawu* 'mother's brother'

Same
(i) *miji* 'husband' | *mata* 'wife'
(ii) *kishiya* 'co-wife'
(iii) *uwargida* 1. 'senior wife'; 2. 'wife'
(iv) *ango* 'bridegroom'; *amarya* 1. 'bride', 2. 'junior wife'
(v) *dan-uba* 'brother (same father, different mother)' | *'yar-uba* 'sister (same father, different mother)'
(vi) 1. *dan'uwa* 1. 'any male person with whom a person has the same father or mother (or both)', 2. 'son of a parent's sister or brother' | *'yar uwa* 1. 'any female person with whom a person has the same father or mother (or both)', 2. 'son of a parent's sister or brother'

[2] The six-system typology, although effective on an operational level, has not remained unchallenged. The main criticism of this classification stems from the fact that it is based on a set of criteria that are partly structural and partly genealogical. For a critique and discussion see Behrens (1984), Kronenfeld (2004, 2006), and Read (2013). The classical typology of kinship structures has been superseded by the algebraic approach (Read and Behrens 1990, Read 2013, Leaf and Read 2021).
[3] Also Ego in anthropological literature.

(vii) *wa* 1. 'elder brother', 2. 'son of father's brother (older than Self)' | 1. *ya* 'elder sister', 2 'daughter of father's brother (older than Self)'
(viii) *ƙane* 1. 'younger brother', 2. 'son of father's brother (younger than Self)' | *ƙanwa* 1. 'younger sister', 2. 'daughter of father's brother (older than Self)'
(ix) *shaƙiƙi* 'full brother' | *shaƙiƙiya* 'full sister'
(x) *taubashi* (also *ɗan-tara-tara*) 'son of one's mother's brother or father's sister (cross-cousin)' | *taubashiya* (also *'yar-tara-tara*) 'daughter of one's mother's brother or father's sister (cross-cousin)'

Descending
(i) *ɗa* 'son' | *'ya* 'daughter'
(ii) *jika* 'grandson' | *jikanya* 'granddaughter'
(iii) *tattaɓa-kunne* 'great grand-son/granddaughter'

In-law
(i) *agola* 'step-son (child of man's wife by her former husband)' | *agoliya* 'step-daughter (child of man's wife by her former husband)'
(ii) *suruki* 'male in-law (e.g. father-in-law, son-in-law, brother-in law) | *suruka* 'female-in-law (e.g., mother-in-law, daughter-in-law, sister-in-law)

It is necessary to point out that although the system, as mentioned above, is (essentially) descriptive, not every combination of distance from Self, gender and type of relationship corresponds to a unique term. In a fully descriptive system (let's leave Morgan's typology aside for a moment and hypothesize a structure that maximizes kin distinctions) in which a distinction is made between siblings born from the same parents, from the same father but with a different mother, and from the same mother but with a different father, we would expect three terms for 'younger brother': a term A designating a sibling younger than Self and sharing the same parents; a term B designating a sibling younger than Self and sharing the same father but not the same mother; and a term C designating a sibling younger than Self and sharing the same mother but not the same father. Instead of three different terms, in Hausa we find a single term, *ƙane*,[4] which group together, within the same generation of Self, different collateral relatives younger than Self. Hence, terms such as *ɗan'uwa*, *wa* e *ƙane* are classificatory, while terms like *shaƙiƙi* 'full brother' or *baffa* 'father's brother' are descriptive.

The terms listed above exclude items such as *bora* 'least loved wife', *iya* 'name given by young children to their mothers', and *mowa* 'favourite wife'. These terms, in fact, are relative and are not fully codifiable in the kinship system. Included in the list but excluded from the analysis is *ango* 'bridegroom' which, although included by Madauci et al. (1968: 25-28) among the terms constituting the *dangi-iyali* complex, does not qualify within the kinship system. *Amarya* will be considered as a kinship term in its acceptation of 'junior wife'.

For the majority of Hausa kinship terms the relative definitions are quite straightforward. Some nouns, however, have two or more acceptations: one core, restricted meaning, and two or more extensions that bleach the specificity of the core kinship term. One case is that of *ɗan'uwa*, a term defined in *Hausa Customs* as "any male person with whom a person has the same father or mother (or both)" (Madauci et al. 1968: 26). Awde (1996) and Newman (2007) assign the term also the notion of cousin, while Kraft and Kirk-Greene extend the definition so as to include "anyone with whom one has any affinity of town, tribe, race, country, creed, trade, color or common humanity" (1994: 223, see also Abraham 1946 below). The full spectrum of meanings assigned to *ɗan'uwa* in lexical descriptions is given below:

[4] From this point on, when discussing individual kinship terms that exist in both masculine and feminine form, I will normally exemplify by using the masculine term, the analysis of the feminine counterpart being entirely specular.

Table 1. Acceptations of *ɗan'uwa*

"Any male person with whom a person has the same father or mother (or both)"	Madauci et al. 1968: 26
"Brother (*strictly* full brother, *but commonly used* for any brother, relative, fellow-country man"	Abraham 1946
1. "Brother (whether full (vide *shaƙiki*); or by same father only; or by same mother only)." 2. "Any relation by blood or marriage."	Bargery 1934
1. Brother 2. Colleague 3. Cousin (in English-Hausa section)	Awde 1996
1. Brother 2. Cousin 3. Relative	Newman 2007

Ɗan'uwa is often used to indicate the relations that Self has within his or her generation. In the present analysis, *ɗan'uwa* will have the acceptations of 'full brother' (i.e. same parents), 'half brother' (i.e. one parent in common), and 'cousin' (i.e. child of a sister or brother of either parent). Note that the more restricted acceptation of *ɗan'uwa* (i.e. 'full brother') overlaps with the specific, non-extended term *shaƙiki*.

Another term with a restricted and an extended meaning is *wa* 'elder brother'. Consider the descriptions of the term provided by Bargery and Abraham:

Table 2. Acceptations of *wa*

1. "Elder brother." 2. "Son (if older than oneself) of father's elder brother." 3. "Any male older than oneself, though not in any way related."	Bargery 1934
1. "Elder brother" 2. "Any male older than oneself, whether related or not"	Abraham 1946

Also in this case, we need to exclude from the analysis the meaning that is extraneous to the kinship system. If, for example, we were to consider the meaning with the maximum extension (as in Abraham's "Any male older than oneself, whether related or not"), we would reduce the intension of the term so drastically (see Section 4) that we could not say anything meaningful about its complexity as a kin term.[5]

3. Theoretical framework and methodological approach

This article intends to analyse lexical data in terms of complexity, and specifically in terms of semantic complexity. It is therefore necessary to define the nature of the linguistic complexity under scrutiny, and on the basis of this to identify the most suitable theoretical framework and methodological approach to achieve the aim.

3.1 A perspective on complexity

Over the last 15 years, the notion of linguistic complexity, i.e. the study of the overall complexity of a language, has interested a significant number of scholars (among others: Miestamo et al. 2008, Sampson et al. 2009, Baerman et al. 2015). The discussion focused and developed on two aspects: 1) the approach to the notion of complexity in absolute or relative terms, and 2) inter-linguistic comparison and the hypothesis that all natural languages are (or are not) equally complex. The absolute

[5] Conversely, an analysis concerned with the metaphorical instantiation of kin terms would also consider terms projected outside the kinship system (cf. Section 8).

approach defines complexity as the number of parts of a system, whereas the relative approach is interested in the relation between cost and difficulty, i.e. the difficulty encountered by an L2 speaker in learning (i.e. processing information about) a certain linguistic phenomenon of an L1 (Kusters 2003, Hawkins 2004). The absolute approach – or at least some of its applications – has made it possible to compare the complexity of different natural languages, with various results: Hockett's hypothesis that all languages are equally complex "since the total grammatical complexity of any language, counting both morphology and syntax, is about the same as that of any other" (Hockett 1958: 180), has for instance been questioned by McWhorter (2001) for creole languages (which are said to be less complex than noncreole languages) and by Miestamo (2008). In particular, Miestamo's criticism (both of Hockett and McWhorter for cases other than creole/non-creole) is directed at the non-quantifiability of the global complexity of a language: the absolute approach to measuring the global complexity of a natural language is necessarily based on the selection of some criteria or grammatical areas and the exclusion of other criteria and grammatical areas, and therefore leads to different results depending on the set of parameters used. What we can really know (and measure) is the "local" complexity of a language, i.e. the complexity of some sub-systems that compose it. The study of local complexities has developed independently of the problem of measuring global complexity (e.g. Lehmann 1978, Maddieson 2005, Bentz et al. 2016, Batic 2020). Attempts to measure individual grammatical areas are the product of different schools (generativism, information theory, typological approach, etc.), just as different are the measurement tools adopted. Although the principle of complexity is based on the number of parts of a system, the calculation must necessarily differ according to the area examined: for example, if for the complexity of a consonantal system it is possible to resort to the quantification of the elements that make up the phonemic inventory (e.g. typological approach, Dryer et al. 2013), for the semantic complexity of a lexical field one must act in a different way, and not for all lexical fields can one make use of the same measurement tools.

3.2 Methodological principles in the measuring of semantic complexity

The representation of a kinship system is based on the terminology of the system and the approach adopted to connect the elements of the system. The relationship between the elements of the system is constructed in a "space" that can be formalized on the basis of either a genealogical or an algebraic approach. The genealogical approach places the kin terms in a space constructed in terms of ascendency-descendency, which allows for a branching and sub-branching representation of the system typical of tree diagram structures. The algebraic approach, on the other hand, is culture-driven and focuses on the way kins refer to each other, that is, kin terms are seen as a product of other kin terms, with the following formalization proposed by Read: "If ego (properly) refers to alter 1 by the kin term L and alter 1 properly refers to alter 2 by the kin term K, then the by the product of K and L, denoted K or L, is meant a kin term (if any) ego properly uses to refer to alter 2" (Read 2013: 3).

The study of the semantic complexity of kin terms, positioning itself in the disciplinary field of lexical semantics, is not so much focused on the representation of the system as such, but rather on the definition of the individual elements (i.e. kin terms) on the basis of defined properties and relations. The study of the absolute complexity (as defined above) of the terms belonging to the lexical field of the kin system is based on the minimum and necessary conditions defining individual kin terms. In this sense, the theoretical-methodological framework consistent with a measurement of complexity values must necessarily isolate properties and formalize relations. Relations that are both appropriate at a formalization level and culture-independent are taken from the genealogical space, which allows each kin term to be analysed as a "sum" of properties. This approach derives from combinatorial semantics and is realized within a feature-matrix system (Katz and Fodor 1963). This system consists in analysing linguistic units (i.e. kin terms) in relation to a finite set of properties (e.g. *x is parent of y*, *x is married to y*, etc.; see Section 5), assigning a truth value to each linguistic unit in relation to the individual properties. The calculation of semantic complexity will thus be based on a truth table, and specifically on the negative dyadic logarithm of the relative number of trues, as illustrated in Section 6.

The advantages of this approach are the following: 1) the method, being culture-independent, is well suited to cross-linguistic comparison; 2) the feature-matrix system allows for the consistent treatment of polysemy instances, as long as they are internal to the kinship terminology; and 3) the semantic complexity values of the kin terms obtained through this approach are applicable to different types of representation, i.e. genealogical or algebraic.[6]

4. Polysemy, extension and intension

The treatment of polysemous words adopted by Lehmann is based on the observation that a polysemous word is characterized by relations (i.e. propositions) that exist as alternatives, that is, they operate in alternative contexts, and therefore they are disjunctively ordered. On the contrary, if two or more relations are operative simultaneously, then they will be considered as propositions conjunctively ordered. The underlying theoretical framework is the feature matrix-based semantic field theory: category membership is determined by sets of sufficient and necessary conditions (Katz and Fodor 1963). According to this theory, polysemy is defined as follows: if a lexical item needs more than one set of sufficient and necessary conditions to account for its meaning, then the item is polysemous. The question that immediately arises concerns the principle governing the role of conjunction and disjunction in assessing the overall complexity of a term.

4.1 Extension and intension

The principle of intention and extension can be exemplified by a series of pairs such as say/whisper, see/inspect, break/shatter: in each of these pairs the first member has a more general meaning, i.e. with greater contextual adaptability, while the second member is more specific, that is, it is used in a more limited range of contexts. Consider for example the Hausa verbs *sha* 'to drink' and *shanye* 'to drink up'. *Sha* is used in a variety of contexts to encode an array of meanings that include, among others, notions such as suffering, inhaling, and eating something pulpy. The verb is mostly productive in idiomatic expressions (*sha wahala* 'suffer' lit. 'to drink problem', *sha kashi* 'to have a hard time' lit. 'to drink shit', *sha tafiya* 'to travel for a long time, to travel far and wide' lit. 'to drink travel'), but has also become grammaticalized as a quantificational verb expressing "multiple/habitual occurrences of an event or situation, where the subject is a volitional agent" (Jaggar and Buba 2009: 244).

(1) *ya* *sha* *zuwa* *nan* (Jaggar and Buba 2009: 144)
3MSG.PFV drink coming here
'He comes here regularly.' (lit. 'He has drunk coming here')

The polysemous nature of *sha* is easily detectable by running a zeugmatic diagnostic test (cf. Cruse 1986: 61-62, Cruse 2000: 108) such as the following:

(2) *na* *sha* *wahala* *da* *mangwaro*
1SG.PFV drink problem and mango
* 'I have drunk problem and mango.'

The oddness of the sentence is due to the activation, within a single context within a single context and through the coordination of two simple noun phrases (*wahala* and *mangwaro*) with the conjunction *da* 'and', of two meanings, neither of which is compatible with the other, of two meanings, neither of which is compatible with the other.[7] In order to function, the two meanings should be activated by two different contexts:

[6] That is, it will be possible to establish correspondences between individual ranges of semantic complexity values and the categories identified by the genealogical (classificatory and descriptive terms) and algebraic (generator terms, nodes, etc.) representations. The application of semantic complexity values to the different representations of the kinship system is beyond the scope of this paper.

[7] The interpretation depends on the nature of *da*. If what followed *wahala* were the comitative preposition *da* 'with' (identical with the conjunction *da* 'and'), then the sentence would acquire a linear meaning: the context would be that of a farmer who encountered difficulties in sowing, planting or cultivating mango trees, or in selling their fruits.

(3) *na* *sha* *wahala* *na* *sha* *mangwaro*
1SG.PFV drink problem 1SG.PFV drink mango
'I have suffered and have eaten a mango.'
(lit. 'I have drunk problem [and] I have eaten a mango')

The *shanye* verb 'to drink up', on the other hand, does not exhibit this flexibility. In addition to its primary meaning, the verb can take on a few other meanings, such as 'to shrink (e.g. cloth after washing)', 'to ignore', and 'to paralyse, to wither (e.g. limb)' (see Bargery 1934, Abraham 1946). *Shanye*, whose usage is essentially restricted to the primary meaning, is a more specialized verb than *sha*. From the point of view of complexity, *sha* is less complex than *shanye* in that its semantic core exists in a multiplicity of meanings, i.e. its greater extension is accompanied by lesser intension. In contrast, *shanye* has a greater intension accompanied by a lesser extension. The principle of extension-intension underlies the calculation of semantic complexity: extension will be formalized as a series of disjuncts, while intension will be translated into a series of conjuncts (see Section 6). Under conditions of comparability, we will notice that polysemic terms (– intension, + extension), such as *ɗan'uwa* 'sibling (general)', have a lower semantic complexity value than non-polysemous terms (+ intension, – extension), e.g. *shaƙiƙi* 'full brother'.

5. Features and relations

Kinship terms are defined according to a set of sufficient and necessary conditions. In the analysis, these conditions correspond to the features usually employed in the description of kin terminologies (or to the "constraints" formulated in Optimality Theory, see Jones 2010) and used, for example, to distinguish sex, distance, grade and generations. The sufficient and necessary conditions needed in order to define any Hausa kinship term are presented in table 3. Each feature has a description expressing a relation or property and a formalization consisting of a simple or binary predicate.

Table 3. Relations and their semantic formalizations

	Description	Formalization	Notes
1	*x is parent of y*	P (x, y)	
2	*x is son of y*	P' (x, y)	converse relation of P (x, y)
3	*x is married to y*	MAR (x, y)	
4	*x is male*	M (x)	
5	*x is female*	F (x)	
6	*x is older than y*	O (x, y)	
7	*x is younger than y*	O' (x, y)	converse relation of O (x, y)

Feature 6 encodes a relation of seniority within the same generation (e.g. elder brother or sister). When used to define affine kin terms such as *uwargida* 'first wife' and *amarya* 'junior wife', the feature works on the culturally (and statistically) grounded assumption that the age of the first wife is greater than the age of the other wives.

6. Semantic calculation

To calculate the semantic complexity of kinship terms within a predicate calculus system, propositions, expressions and operators are used. A proposition is a molecular statement consisting of one relation; an expression is constituted by at least two propositions or by a series of propositions that can be either conjuncts, if separated by the conjunction $\wedge$, or disjuncts, if separated by the disjunction $\vee$.

6.1 Formulas

The semantic complexity is expressed by a coefficient *c* calculated as the negative dyadic logarithm of the relative number of trues. The general formula proposed by Lehmann is based on the amount of semantic information *inf* postulated by Carnap and Bar-Hillel (1952).

General formula:

$$c = log_2 \frac{1}{x/2^n} = n - log_2 x$$

where c is the complexity coefficient and n the number of conjoined propositions

Formula #1 – Conjoined propositions: $p_1 \wedge p_2 \wedge p_3 \wedge \ldots p_n$
$c = n$
for any n, x = 1

Most kinship terms are encoded by series of conjoined propositions. The term *uba* 'father', for example will consist of two statements: a proposition specifying the relation P (x, y) (i.e. x is parent of y) and a second proposition M (x) (i.e. x is male) specifying the gender:

x *uba* y   P (x, y) ∧ M (x)

Thus, applying formula #1, we will obtain $c = 2$.

Formula #2 – Disjoined propositions: p ∨ q
$c = n - log_2 x$
$x = 2^n - 1$

Polysemous words are encoded through disjoined propositions. A simple disjunction consists of two propositions (p ∨ q).

Formula #3 – Complex disjoined expressions: $(p_1 \wedge p_2 \wedge p3 \wedge \ldots p_m) \vee (q_1 \wedge q_2 \wedge q_3 \wedge \ldots q_n)$
$c = m + n - log_2(2^m + 2^n - 1)$
$x = 2^n + 2^m - 1$

Within a kinship system, polysemous words consist of complex disjoined expressions: each disjuncts is constituted by two or more conjoined propositions. Consider for example the case of *babba*, a term that has the meaning of 'elder, senior' and is also used in addressing grandfathers. Let us assume, just for explanatory purposes, that *babba* is a polysemous word with two meanings: 'elder' and 'grandfather'. The 'elder' meaning will need the following series of conjunctions: P (z, y) ∧ O (x, z)[8] ∧ M (x). 'Grandfather', on the other hand, will be encoded through the following series of propositions: P (x, z) ∧ P (z, y) ∧ M (x). The full disjunction will read as follows:

(P (z, y) ∧ O (x, z) ∧ M (x)) ∨ (P (x, z) ∧ P (z, y) ∧ M (x))

Since there are propositions that appear in both terms, it will be necessary to extract them by applying the distributive law:

(P (z, y) ∧ M (x)) ∧ (O (x, z) ∨ P (x, z))

The c of each conjunct can be calculated separately. Therefore, we will add the c of P (z, y) ∧ M (x) to the c of the disjunction O (x, z) ∨ P (x, z). The c of the disjunction is obtained by applying formula #2. The c of the entire expression will be:

$c = 2 + (2 - log_2 3) = 2.42$

[8] As stated in Section 5, the relation O 'older than' is not absolute, but relative to a specific generation, i.e. both terms of the binary predicate are understood as belonging to the same generation.

6.2 Conventions

The following conventions have been adopted:

1. x always occupies the first position of the relation, *y* the second;
2. $z_1, z_2, z_3 \dots z_n$ can occur in any position;
3. in the genealogical space, *y* stands for Self;
4. the apostrophe indicates a converse relation and is counted as a proposition, hence a proposition like $P'(x, y) \wedge M(x)$ will give $c = 3$;
5. the semantic analysis requires the sex be specified for both masculine and feminine forms (see discussion below);
6. for children with a common parent, it is necessary to state that they are not identical, i.e. $x \neq y$;
7. distributive and absorption rules are applied when identical propositions are present in two or more disjunctions.

In his analysis of kinship terms, Lehmann uses a relation specifying the gender only for female terms, namely F (x). Male terms, on the other hand, never include a gender-specific relation M (x), except for the term *marido* 'husband'. Lehmann's explanation is that "the male terms [...] cannot be specified by M (x) because they are not necessarily male" (1978: 8). This is based on the fact that in Portuguese the unmarked gender is masculine, e.g. to refer indistinctly to uncles and aunts a speaker would normally use the masculine plural form (*tios*) and not the feminine one (*tias*). In Hausa, however, the phenomenon is different. The gender distinction, which is productive in the singular, is neutralized in the plural. The masculine singular and plural thus fall into two distinct grammatical categories (Newman 2000: 200). The only terms that will not include a gender-specific relation are the epicene *kaka* 'grandfather, grandmother' and *tattaɓa-kunne* 'grandchild', two nouns whose gender depends on the gender of the referent.

7. Complexity values

The kinship terms are ordered as follows: ascending, ascending-lateral, same, descending, and in-law. Masculine and epicene terms are indicated on the left, feminine terms on the right. The value of the complexity coefficient *c* is indicated in square brackets to the left of the kinship term.

Ascending

[2]	x **uba** y	$P(x, y)$ $\wedge\ M(x)$	[2]	x **uwa** y	$P(x, y)$ $\wedge\ F(x)$

[2]	x **mahaifi** y	$P(x, y)$ $\wedge\ M(x)$	[2]	x **mahaifiya** y	$P(x, y)$ $\wedge\ F(x)$

[2]	x **kaka** y	$P(x, z)$ $\wedge\ P(z, y)$			

Ascending-lateral

[7]	x **baffa** y	$P(z_1, y)$ $\wedge\ P(z_2, z_1)$ $\wedge\ P'(x, z_2)$ $\wedge\ M(z_1)$ $\wedge\ x \neq z_1$ $\wedge\ M(x)$	[7]	x **gwaggo** y	$P(z, y)$ $\wedge\ P(z_2, z_1)$ $\wedge\ P'(x, z_2)$ $\wedge\ M(z_1)$ $\wedge\ x \neq z_1$ $\wedge\ F(x)$

[7]	x **rafani** y	$P(z_1, y)$	[7]	x **inna** y	$P(z, y)$
		$\wedge\ P(z_2, z_1)$			$\wedge\ P(z_2, z_1)$
		$\wedge\ P'(x, z_2)$			$\wedge\ P'(x, z_2)$
		$\wedge\ F(z_1)$			$\wedge\ F(z_1)$
		$\wedge\ x \neq z_1$			$\wedge\ x \neq z_1$
		$\wedge\ M(x)$			$\wedge\ F(x)$

[7]	x **kawu** y	$P(z_1, y)$
		$\wedge\ P(z_2, z_1)$
		$\wedge\ P'(x, z_2)$
		$\wedge\ x \neq z_1)$
		$\wedge\ F(z_1)$
		$\wedge\ M(x)$

Same

[2]	x **miji** y	$MAR(x, y)$	[2]	x **mata** y	$MAR(x, y)$
		$\wedge\ M(x)$			$\wedge\ F(x)$

[3]	x **kishiya** y	$MAR(x, y)$
		$\wedge\ MAR(z, y)$
		$\wedge\ F(x)$

[4]	x **uwargida** y	$MAR(x, y)$
		$\wedge\ MAR(z, y)$
		$\wedge\ O(x, z)$
		$\wedge\ F(x)$

[5]	x **amarya** y	$MAR(x, y)$
		$\wedge\ MAR(y, z)$
		$\wedge\ O'(x, z)$
		$\wedge\ F(x)$

[8]	x **shaƙiƙi** y	$P'(x, z_1)$	[8]	x **shaƙiƙiya** y	$P'(x, z_1)$
		$\wedge\ P(z_1, y)$			$\wedge\ P(z_1, y)$
		$\wedge\ P'(x, z_2)$			$\wedge\ P'(x, z_2)$
		$\wedge\ P(z_2, y)$			$\wedge\ P(z_2, y)$
		$\wedge\ x \neq y$			$\wedge\ x \neq y$
		$\wedge\ M(x)$			$\wedge\ F(x)$

[6.5]	x **ƙane** y[9]	$(P(x_1, y)$	[6.5]	x **ƙanwa** y	$(P(x_1, y)$
		$\wedge\ O'(x, y)$			$\wedge\ O'(x, y)$
		$\wedge\ M(x))$			$\wedge\ F(x))$
		$\wedge\ ((P'(x, z_1)$			$\wedge\ ((P'(x, z_1)$
		$\wedge\ x \neq y)$			$\wedge\ x \neq y)$
		$\vee\ (P(z_2, z_1)$			$\vee\ (P(z_2, z_1)$
		$\wedge\ P(z_2, z_3)$			$\wedge\ P(z_2, z_3)$
		$\wedge\ P'(x, z_3)$			$\wedge\ P'(x, z_3)$
		$\wedge\ z1 \neq z_3))$			$\wedge\ z1 \neq z_3))$

[9] $(P(z_1, y) \wedge O'(x, y) \wedge M(x)) \wedge ((P'(x, z_1) \wedge x \neq y) \vee (P(z_2, z_1) \wedge P(z_2, z_3) \wedge P'(x, z_3) \wedge z1 \neq z_3))$
Applying formulas #1 and #3:
$c = 4 + (3 + 4 - log_2(2^3 + 2^4 - 1)) = 6.5$

[5.7] x **wa** y[10]	$(P (z_1, y)$	[5.7] x **ya** y	$(P (z_1, y)$
	$\wedge\ O (x, y)$		$\wedge\ O (x, y)$
	$\wedge\ M (x))$		$\wedge\ F (x))$
	$\wedge\ ((P' (x, z_1)$		$\wedge\ ((P' (x, z_1)$
	$\wedge\ x \neq y)$		$\wedge\ x \neq y)$
	$\vee\ (P (z_2, z_1)$		$\vee\ (P (z_2, z_1)$
	$\wedge\ (P (z_2, z_3)$		$\wedge\ (P (z_2, z_3)$
	$\wedge\ z1 \neq z_3$		$\wedge\ z1 \neq z_3$
	$\wedge\ P' (x, z_3)))$		$\wedge\ P' (x, z_3)))$

[8] x **ɗan-uba** y	$P' (x, z_1)$	[8] x **'yar-uba** y	$P' (x, z_1)$
	$\wedge\ P (z_1, y)$		$\wedge\ P (z_1, y)$
	$\wedge\ P (z_2, y)$		$\wedge\ P (z_2, y)$
	$\wedge\ P' (x, z_3)$		$\wedge\ P' (x, z_3)$
	$\wedge\ M (z_1)$		$\wedge\ M (z_1)$
	$\wedge\ M (x)$		$\wedge\ F (x)$

[4.75] x **ɗan'uwa** y[11]	$(P' (x, z_1)$	[4.75] x **'yar uwa** y	$(P' (x, z_1)$
	$\wedge\ ((P (z_1, y)$		$\wedge\ ((P (z_1, y)$
	$\wedge\ x \neq y)$		$\wedge\ x \neq y)$
	$\vee\ (P (z_2, z_1)$		$\vee\ (P (z_2, z_1)$
	$\wedge\ P (z_2, z_3)$		$\wedge\ P (z_2, z_3)$
	$\wedge\ P (z_3, y)$		$\wedge\ P (z_3, y)$
	$\wedge\ z_1 \neq z_3))$		$\wedge\ z_3 \neq y)))$
	$\wedge\ M (x)$		$\wedge\ F (x)$

[8.2] x **taubashi** y[12]	$(P (z_1, y)$	[8.2]x **taubashiya** y	$(P (z_1, y)$
	$\wedge\ P (z_2, z_1)$		$\wedge\ P (z_2, z_1)$
	$\wedge\ P (z_2, z_3)$		$\wedge\ P (z_2, z_3)$
	$\wedge\ P' (x, z_3)$		$\wedge\ P' (x, z_3)$
	$\wedge\ z_1 \neq z_3$		$\wedge\ z_1 \neq z_3$
	$\wedge\ M (x))$		$\wedge\ F (x))$
	$\wedge\ ((F (z_1)$		$\wedge\ ((F (z_1)$
	$\wedge\ M (z_3))$		$\wedge\ M (z_3))$
	$\vee\ (M (z_1)$		$\vee\ (M (z_1)$
	$\wedge\ F (z_3)))$		$\wedge\ F (z_3)))$

Descending

[3] x **ɗa** y	$P' (x, y)$	[3] x **'ya** y	$P' (x, y)$
	$\wedge\ M (x)$		$\wedge\ F (x)$
[5] x **jika** y	$P' (x, z)$	[5] x **jikanya** y	$P' (x, z)$
	$\wedge\ P' (z, y)$		$\wedge\ P' (z, y)$
	$\wedge\ M (x)$		$\wedge\ F (x)$

[10] $(P (z_1, y) \wedge M (x) \wedge O (x, y)) \wedge ((P' (x, z_1) \wedge x \neq y) \vee (P (z_2, z1) \wedge P (z_2, z_3) \wedge z_1 \neq z_3 \wedge P' (x, z_3)))$
Applying formulas #1 and #3:
$c = 3 + (3 + 5 - log_2(2^3 + 2^5 - 1)) = 5.7$

[11] $P' (x, z_1) \wedge ((P (z_1, y) \wedge x \neq y) \vee (P (z_2, z_1) \wedge P (z_2, z_3) \wedge P (z_3, y) \wedge z_1 \neq z_3)) \wedge M (x)$
Applying formulas #1 and #3:
$c = 2 + (2 + 4 - log_2(2^2 + 2^4 - 1)) + 1 = 4.75$

[12] $(P (z_1, y) \wedge P (z_2, z_1) \wedge P (z_2, z_3) \wedge P' (x, z_3) \wedge z_1 \neq z_3 \wedge M (x)) \wedge ((F (z_1) \wedge M (z_3)) \vee (M (z_1) \wedge F (z_3)))$
Applying formulas #1 and #3:
$c = 7 + (2 + 2 - log_2(2^2 + 2^2 - 1)) = 8.2$

[5]	x **tattaɓa-kunne** y		
			P' (x, z_1)
		∧	P (z_2, z_1)
		∧	P' (z_2, y)

In-law

[7]	x **agola** y		P' (x, z_1)	[7]	x **agoliya** y		P' (x, z_1)
		∧	MAR (z_1, y)			∧	MAR (z_1, y)
		∧	P' (x, z_2)			∧	P' (x, z_2)
		∧	F (z_1)			∧	F (z_1)
		∧	M (x)			∧	F (x)

(father-in law / mother-in-law)

[4]	x **suruki** y		P (z_1, z_2)	[4]	**suruka**		P (z_1, z_2)
		∧	MAR (x, z_1)			∧	MAR (x, z_1)
		∧	MAR (z_2, y)			∧	MAR (z_2, y)
		∧	M (x)			∧	F (x)

(son-in-law / daughter-in-law)

[4]	x **suruki** y		P' (z, y)	[4]	**suruka**		P' (z, y)
		∧	MAR (x, z)			∧	MAR (x, z_2)
		∧	M (x)			∧	F (x)

(brother-in-law / sister-in-law)

[4]	x **suruki** y		P (z_1, y)	[4]	**suruka**		P (z_1, y)
		∧	P (z_1, z_2)			∧	P (z_1, z_2)
		∧	MAR (x, z_2)			∧	MAR (x, z_2)
		∧	M (x)			∧	F (x)

8. Discussion

The results given by the calculation of the complexity coefficient are consistent with the distribution and numerical consistency of the terms in the different generations. In the generation labelled "same", for example, we find a rather high number of terms belonging to the category "siblings". Specifically, it is in this generational band that we identify terms that classical kinship science defines as descriptive. Descriptive terms concern a restricted sub-category of items; they are those with the highest complexity coefficient: highly specific terms such as *shaƙiƙi* 'full brother' (c = 8), *taubashi* 'cross-cousin' (c =8.2), and *ɗan-uba* 'half-brother (same father, different mother)' (c = 8) have complexity coefficients equal to or greater than 8. In the same generational band is also the classificatory term *ɗan'uwa* 'brother, cousin', whose c is less than 5 (adopting Kraft and Kirk-Greene's definition above and extending the relation beyond the *dangi* network, the c of the lexeme would tend even to a lower value). Terms such as *wa* (c = 5.7) and *ƙane* (c = 6.5), while not presenting a formalization too distant from *ɗan'uwa*, owe their higher coefficient to the seniority relation that characterizes them.

Sauraki 'in-law (male)' is not a proper kinship term, but rather a general attribute that identifies specific kinship roles that are inferable only from the context. For *sauraki*, in its general sense, it is not possible to construct a genealogical space in the same way as has been done for the other terms. The complexity calculation is conducted on the realizations of *sauraki* within the genealogical space (father-in-law, brother-in-law, son-in-law). Note how the three realizations of the term have a c equal to 4. A highly complex term belonging to the first descending generation is *agola* (c = 7).

The kinship terms defined in the previous section are formulated from a set of generating relations. The underlying representation of the terminology is genealogical: x and y are always defined on the basis of their connection to a common ancestor (in the case of non-affine relations). For example, to define <x *rafani* y>, read x is *rafani* of y, the following relations are specified: z_1 is parent of y, z_2 is parent of z_1, and x is child of z_2 (in addition to the relations specifying the difference of x and z_1 and the gender of x). The structure of the terminology is thus anchored to the classification of a genealogical space. One aspect of this type of structure is that the kinship space is defined a priori: it is possible to define the relationship between x and y, i.e. the term that a member of the *dangi-iyali* corresponding to y (Self) uses to refer to another member of the *dangi-iyali* corresponding to x, because the position of x within the kinship network is already known. This analytical approach is useful insofar as the objective is to measure the amount of semantic information independently of the cultural context, namely independently of the construction of the kinship terminology as a conceptual network operating in a given culture. As observed and discussed by several scholars (among others, Sahlins 1962, Good 1981, and notably Read and Behrens 1990, Read 2006, 2013), a member of a given society calculates his or her kin term relation to another person via the kin term relation each has to a third person, without resorting to genealogical reconstruction. The algebraic approach differs from the genealogy-based structure in two respects: 1) it defines the terms of kinship relations from generating kin terms, and 2) it is culturally grounded and views kin terms as forming a system of symbols (Read 2006, 2013).

However, not even the algebraic approach, like the genealogical one, can represent the cultural characterization associated with different kinship terms, such as the hierarchy of the members of the system or the rights and duties inherent in individual roles. Consider the role of *inna*, the sister of Self's mother: in Hausa society, it is *inna* (and not *goggo*) who acts as mother when Self's mother dies, but this inherent task in the role of *inna* is not encoded in the system's representation of kinship terms. More generally, genealogical and algebraic analyses exclude the metaphorical or honorary or extended usage of kin terms.

Conclusion

The study analysed the semantic complexity of Hausa kinship terms using a calculation system built on the principle of intention and extension and on the quantification of the semantic information operated on the propositional formalization of the terms. The methodology adopted is applicable to semantic fields whose items are transposable in a feature-matrix system. Semantic complexity analysis confirms the descriptive-classificatory nature of Hausa kinship terminology. High value terms are descriptive, while lower value terms are classificatory. Moreover, the article shows how the analysis of the semantic complexity – or semantics in general – of kinship terms cannot rely entirely on the lexical descriptions we possess. The uneven treatment of kinship terms in the dictionaries of African languages (especially in the more compact and modern versions) denotes a certain haste in treating English kinship terms as meta-terms, when entries for kinship terms often need a description and not just a gloss. This is particularly true for non-generating terms, such as collateral kinship terms.

References

Abraham, R.C., 1946. *Dictionary of the Hausa language*. London: London University Press.

Awde, N., 1996. *Hausa-English, English-Hausa dictionary*. New York: Hippocrene Books.

Baerman, M. et al., 2015. *Understanding and measuring morphological complexity*. Oxford: Oxford University Press.

Bargery, G.P., 1934. *A Hausa-English dictionary and English-Hausa vocabulary*. London: Oxford University Press.

Batic, G.C., 2020. Measuring phonological complexity in West African languages. In N. Pawlak and I. Will, eds. *West African languages. Theory and communication*. Warsaw: University of Warsaw, pp. 23-45.

Behrens, C., 1984. *Shibipo ecology and economy*. PhD dissertation, Los Angeles: UCLA.

Bentz, C. et al. 2016. A comparison between morphological complexity measures: Typological data vs. language corpora. *Proceedings of the Workshop on Computational Linguistics for Linguistic Complexity* (Osaka, Japan, December 11-17 2016), pp. 143-153.

Bernard, H.R., 2011. *Research methods in anthropology: Qualitative and quantitative approaches*. Lanham: AltaMira Press.

Carnap, R. and Bar-Hillel, Y., 1952. An outline of a theory of semantic information. *Technical Report*, no. 247. Cambridge, Mass.: MIT Research Laboratory of Electronics.

Cruse, A.D., 1986. *Lexical semantics*. Cambridge: Cambridge University Press.

Cruse, A.D., 2000. *Meaning in languages*. Oxford: Oxford University Press.

Dryer, M.S. and Haspelmath, M., eds. 2013. *The world atlas of language structures online*, Leipzig: Max Planck Institute for Evolutionary Anthropology. [Accessed on 03 June 2022] Available at http://wals.info

Good, A., 1981. Prescription, preference and practice: Marriage patterns among the Kondaiyankottai Maravar of South India. *Man* (N.S.), no. 16, pp. 108-129.

Hawkins, J.A., 2004. *Efficiency and complexity in grammars*. Oxford: Oxford University Press.

Hockett, C.F., 1958. *A course in modern linguistics*. New York NY: Macmillan.

Jaggar, P.J. and Buba, M., 2009. Metaphorical extension of 'eat' => [OVERCOME] and 'drink' => [UNDERGO] in Hausa. In: J. Newman, ed. *The Linguistics of Eating and Drinking*. Amsterdam and Philadelphia: John Benjamins, pp. 229-251.

Jones, D., 2010. Human kinship, from conceptual structure to grammar. *Behavioral and Brain Sciences*, no. 33, pp. 367-416.

Katz, J. and Fodor, J.A., 1963. The structure of semantic theory. *Language*, no. 39, pp. 170-210.

Kirchhoff, P., 1932. Verwandtschaftsbezeichnungen und Verwandtenheirat. *Zeitschrift für Ethnologie*, no. 64, pp. 41-71.

Kraft, C.H. and Kirk-Greene, A., 1994. *Hausa*. Lincolnwood, Il: NTC Publishing Group.

Kronenfeld, D.B., 2004. Definitions of cross versus parallel: Implications for a new typology (an appreciation of A. Kimball Romney). *Cross-Cultural Research*, no. 38, pp. 249-269.

Kronenfeld, D.B., 2006. Issues in the classification of kinship terminologies. Toward a new typology. *Anthropos*, vol. 101, no. 2, pp. 203-219.

Kusters, W., 2003. *Linguistic complexity, the influence of social change on verbal inflection*, Ph.D. Diss., University of Leiden, Utrecht: LOT.

Leaf, M.J. and Dwight W.R., 2021. *Introduction to the science of kinship*. Lanham: Lexington Books.

Lehmann, C., 1978. On measuring semantic complexity. A contribution to a rapprochement of semantics and statistical linguistics. *Georgetown University papers in languages and linguistics*, no. 14, pp. 83-120.

Lowie, R. H., 1928. A note on relationship terminologies. *American Anthropologist*, no. 30, pp. 263-267. https://doi.org/10.1525/aa.1928.30.2.02a00060

Ma Newman, R., 1990. *An English Hausa-Dictionary*. New Haven & London: Yale University Press.

Madauci, I., Isa, Y., and Daura, B., 1968. *Hausa Customs*. Zaria: Northern Nigeria Publishing Company.

Maddieson, I., 2005. Issues of phonological complexity: Statistical analysis of the relationship between syllable structures, segment inventories and tone contrasts. In *UC Berkeley Phonology Lab Annual Report*, Berkeley: University of California, pp. 269-268.

McWhorter, J., 2001. The world's simplest grammars are creole grammars. *Linguistic Typology*, no. 5, pp. 125-166.

Miestamo, M., Sinnemäki, K., and Karlsson, F., 2008. *Language complexity: Typology, contact, change*. Amsterdam: John Benjamins.

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

Murdock, G.P., 1949. *Social Structure*. New York: Macmillan.

Newman, P., 2000. *The Hausa language. An encyclopedic reference grammar*. New Haven & London: Yale University Press.

Newman, P., 2007. *A Hausa-English dictionary*. New Haven & London: Yale University Press.

Read, D.W. 2013. A new approach to forming a typology of kinship terminology systems: From Morgan and Murdock to the present. *Structure and Dynamics*, vol. 6, no. 1, pp. 1-28.

Read, D.W. 2006. Kinship algebra expert system. A software implementation of a cultural theory. *Social Science Computer Review*, vol. 24, no. 1, pp. 43-67.

Read, D.W. and Behrens, C., 1990. KAES: An expert system for the algebraic analysis of kinship terminologies. *Journal of Quantitative Anthropology*, no. 2, pp. 353-393.

Sahlins, M.D., 1962. *Moala: Culture and nature on a Fijian island*. Englewood Cliffs, NJ: Prentice Hall.

Sampson, G., Gil D., and Trudgill, P., 2009. *Language complexity as an evolving variable*. Oxford: Oxford University Press.

Spier, L., 1925. The Distribution of kinship systems in North America. *University of Washington Publications in Anthropology* no. 1, pp. 69-88.
Zwicky, A. and Sadock, J., 1975. Ambiguity tests and how to fail them. In: J. Kimball, ed. *Syntax and Semantics 4*. New York: Academic Press, pp. 1-36.