Acquiring the lexicon and grammar of universal kinship

Joe Blythe
Macquarie University
joe.blythe@mq.edu.au

Jeremiah Tunmuck
Yek Yederr

Alice Mitchell
University of Cologne
alice.mitchell@uni-koeln.de

Péter Rácz
Central European University
raczp@ceu.edu
Acquiring the lexicon and grammar of universal kinship
This paper investigates how children learn an infinitely expanding ‘universal’ system of classificatory kinship terms. We report on a series of experiments designed to elicit acquisitional data on (i) nominal kinterms and (ii) sibling inflected polysynthetic morphology in the Australian language Murrinhpatha. Photographs of the participants’ own relatives are used as stimuli to assess knowledge of kinterms, kin-based grammatical contrasts, and kinship principles, across different age groups. The results show that genealogically distant kin are more difficult to classify than close kin; that children’s comprehension and production of kinterms is streamlined by abstract merging principles; and that sibling-inflection is learned in tandem with number and person marking in the verbal morphology, although it is not fully mastered until mid to late childhood. We discuss how the unlimited nature of Australian kinship systems present unusual challenges to the language learner, but suggest that, as everywhere, patterns of language acquisition are closely intertwined with children’s experience of their sociocultural environment.*

*Acknowledgements:
Many thanks for the feedback and questions we have received from two anonymous reviewers, and the editorial team at Language. We’re also grateful for feedback received on presentations of this work at CAKTAM (Children’s Acquisition of Kinship Knowledge: Theory and Method) in Bristol, Kinship Cognition and Practice at the Max Planck Institute for Social Anthropology in Halle and at the 2019 conference of the Australian Linguistic Society. Many thanks to Lucinda Davidson, John Mansfield, Bill Forshaw, Fiona Jordan, Linda Barwick, Rachel Nordlinger, Kinngirri Carmelita Perdjert, Mawurt Ernest Perdjert, Loraine Tunmuck and Leo Weardon for invaluable feedback, advice and support. The research was funded by the Australian Research Council (DE130100399, DP110100961).

*Keywords: language acquisition, kinterms, kintax, polysynthetic languages, semantic categories, morphology*
1. INTRODUCTION.

In order to talk about the people we meet we need to learn who they are and how they fit within the various social networks we move in. For the child, this includes learning which people may be referred to as ‘family’ and how. In large urban industrialized societies the number of people considered kin is relatively small. Children generally encounter these family members within the first few years of their lives. In doing so they acquire enough knowledge of how kinterms are used to know who is being spoken about and how to address these relatives appropriately, although adult-like conceptualisation of kinterms, particularly the ability to take different perspectives on kinship relations, may take longer to achieve (Piaget 1928). For many children around the world, kinship is a relatively restricted domain, with but a handful of uncles and aunts, parents, grandparents, and a few cousins to learn about; no-one else is considered ‘family’. This situation is drastically different, however, if you live in a small community where virtually everyone you meet can be referred to with a kinterm. In this kind of environment, the system of kinship relations becomes inherently more complex and the magnitude of the acquisition task is exceedingly greater. And if abstract kinship concepts like siblinghood or generational harmony are core grammatical features of the language you speak, then questions of ‘relatedness’ concern all people, all the time, and not merely one’s immediate family on specific occasions. This posits an unusual problem for the language learner. This paper investigates children’s knowledge of kinterms and kin-based grammar in such a community, namely Wadeye in Australia’s Northern Territory, where Murrinhpatha is the dominant language.

In his seminal volume on kinship, Lewis Henry Morgan (1871) distinguished between DESCRIPTIVE kinterms like the English term mother, which specify a unique type of relationship (a person’s female parent), and CLASSIFICATORY terms like uncle which denote a collection of possible relationships (MB, FB, MZH, FZH), thus allowing a greater number of possible referents. He also classified certain terminological systems (e.g. Sudanese) as highly descriptive and others (e.g. Hawaiian) as highly classificatory. Australian ‘universal’ kinship systems like the Murrinhpatha system could be considered extremely classificatory because they allow every individual in the social universe to be allocated a kinterm. Thus, the Murrinhpatha term for ‘mother’ kale is also used for other close consanguineal kin (MZ), affinal kin (FW), and yet other relatives of greater genealogical distance (MMD, MMZD, FFSW, FFBSW, MMMDD, MMMZDD, etc.). The categories and rules enabling kinship systems to expand indefinitely make
acquiring such systems a cognitively demanding and protracted process that likely extends into adulthood. What is more, Murrinhpatha speakers make kin-based morphosyntactic distinctions, encoding siblinghood in the verbal paradigm. Children’s acquisition of inflectional morphology, then, depends on an understanding of kinship relations.

While there has been long-standing interest within developmental psychology and sociocultural anthropology in children’s acquisition of restricted sets of kinterms, very little attention has been given to the acquisition of classificatory systems of kinship terms. And while the expression of kinship relations in grammar has been a long-standing interest for anthropological linguistics (Hale 1966; Schebeck 1973; Hercus & White 1973; Koch 1982; Alpher 1982; Dahl & Koptjevskaja-Tamm 2001; Evans 2003; Blythe 2013; Keen 2013), it has been of little concern to the field of child language acquisition. Core kin-based morphosyntax has never been comprehensively investigated from an acquisitional perspective. Here we report on the results of two experiments designed to elucidate Murrinhpatha children’s understanding of kinterms and kin-based grammatical distinctions. Unlike earlier work on children’s understandings of kinterms, which generally tested children’s ability to define kinterms, our tasks were personalised for each child and used photographs to test knowledge of real-world kin relations.

In the following section we ground our research in the context of prior acquisitional studies of kinship concepts and social categories. Within the broader context of classificatory kinship systems we then provide some background about the Murrinhpatha language and its speakers (in §3), demonstrating how kinship concepts cross-cut the nominal lexicon and verbal morphology. From there we provide the rationale behind the design of the research (in §4), before presenting results from our studies of nominal kinterms (in §5) and verbal morphosyntax (in §6). Finally (in §7) we suggest mechanisms through which classificatory kinship categories begin to develop as children learn to practically apply abstract merging principles (see 3.1) in their daily lives.

Our results demonstrate that language and the social context of its learning are deeply intertwined in Wadeye. Our innovative methodological approach brings typologically rare linguistic data to the field of language acquisition, which is important because almost nothing is known about how diverse features of grammar impact on the social and cognitive development of children. Our finding that sibling inflected morphosyntax is not more difficult to learn, per se, than gender and number contrasts, was unexpected. This discovery is significant for linguistics in
general, and language acquisition in particular, as it suggests that developmental advances
normally attributed to a quite restricted set of semantic domains (e.g. gender, number, person)
may be more generalised than was previously thought. Furthermore, our methodological
distinction between **GENEOLOGICAL DISTANCE** and **KINSHIP DISTANCE** (see 5.3) has revealed that
abstract merging principles (see 3.1) appear to have an ontological reality for children, in that
they begin to expand classificatory kinship categories indefinitely as kinship mergers are
incrementally acquired. For both sociocultural anthropology and anthropological linguistics, this
makes an important and practical contribution toward the theoretical underpinnings of kinterm
semantics.

2. **CHILDREN’S ACQUISITION OF KINSHIP TERMS.**

Kinship terminology has been of enduring fascination to anthropologists (e.g. Radcliffe-Brown
1930/1931; Lounsbury 1964; Kronenfeld 2008), yet little anthropological research has explored
the cultural transmission of these terminological systems. How children learn kinterms also
presents an interesting problem for linguists and psychologists: these terms denote categories that
are (a) highly culture-specific, and (b) relational in nature (Gentner & Kurtz 1995), that is, they
are reckoned from the perspective of a shifting ‘ego’. At what age can children correctly identify
types of kin? Are some categories harder to learn than others? What can acquisitional patterns
tell us about the meaning of kinterms? In this section we provide an overview of earlier work on
children’s acquisition of kinship concepts to provide some comparative context for the
Murrinhpatha study.

Most acquisitional research on kinship cognition within developmental psychology builds
on the work of Piaget (1928). Piaget was interested in the logical properties of kinterms and used
them to test children’s perspective-taking abilities. Piaget (1928) asked children about their own
brothers and sisters, and the relations between them, as well as eliciting definitions of the words
‘brother’/‘sister’. He showed that children under the age of about nine have trouble describing
themselves as a brother or sister to someone else (at least in the context of his questionnaire) and
used this as evidence for the cognitive egocentricity of the young child. Most of the studies
replicating Piaget’s methods recognize four developmental stages in conceptualizing kin
relations: a **pre-categorical** stage in which a form like *brother* is understood as equivalent to a
name like *John*; a **categorical** stage in which the child overgeneralizes a semantic feature of a
kinterm like age or gender, thus extending a term like brother to ‘boys’ in general; a relational stage in which kinterms are reckoned egocentrically from the child’s perspective only, such that the relationship is understood as unidirectional; and a reciprocal stage in which the child understands a term like father to also implicate an inverse (or altercentric) relationship (son/daughter) (Haviland & Clark 1974; LeVine & Price-Williams 1974; Greenfield & Childs 1977; Price-Williams et al. 1977). These studies, conducted in a variety of languages, have largely confirmed that children’s definitions of kinterms progress through these different stages and that, at least in their definitions, children can successfully take alternative perspectives on kinship relations by around the age of eight.

In K. Danziger’s (1957) study of Anglo-Australian children’s kinship concepts, he observed that children’s definitions of certain kinterms were more complex than other definitions, suggesting that some kinterms are easier to learn than others. Haviland and Clark (1974) explored this observation further by eliciting definitions of fifteen kinterms which varied in their compositional semantics. They argue that the order in which children acquire kinterms is determined by semantic complexity, thus terms with fewer relational components, like father ([X Parent of Y][Male X]), were acquired before terms with more relational components, like nephew ([X Child of A][A Sibling of Y][Male X]). E. Danziger (1993) makes use of the idea that acquisitional patterns reflect semantic complexity to pit the predictions of alternative componential analyses of two Mopan kinterms against each other. She found that Mayan children’s (age range 7–14) progression through the Piagetian stages of acquisition aligned better with a categorical/monosemic analysis based on common features than with a genealogical/polysemic analysis based on typical referents (E. Danziger 1993; or with a homonymic analysis, 2001). On the other hand, in their study with Zinacantec Maya children, Greenfield and Childs (1977) found that, of three different analyses of semantic complexity, none correctly predicted the order of acquisition of sibling terms.

Some researchers (e.g. Benson & Anglin 1987; Bavin 1991) have proposed that a child’s individual experience with kin is a crucial factor for explaining the order in which they master kinterms. The socio-cultural significance of kinship and types of kin within a given community may also influence learning. In the only truly comparative research that we know of, Ragnarsdóttir (1997; 1999) showed the importance of cultural and demographic, as opposed to morphological and semantic factors in learning about kinship. In a study comparing Danish and
Icelandic children aged 4–8, Ragnarsdóttir (1999) found that Icelandic children were about a year ahead of their Danish counterparts in reasoning about kinship. She explains this difference partly in terms of the cultural preoccupation with genealogy and kinship in Iceland, as well as the small and homogeneous population. She also points out that the linguistic form of kinterms does not appear to aid acquisition: unlike Icelandic, certain Danish kinterms transparently distinguish patrilineal and matrilineal descent, yet the Danish children performed less well on questions concerning the father and mother of their parents.

Much of the early research on children’s kinship concepts is marred by its reliance on definition-based tasks. Hirschfeld (1989) argues that definitions only reveal children’s metaconceptual, rather than conceptual, knowledge. Haviland and Clark (1974) note that definition tasks only give a partial insight into what a child knows about a particular kinterm. Another shortcoming of definition tasks is (lack of) cross-cultural validity: providing definitions abstracted from context may be a marginal or non-existent metalinguistic practice in some communities. Other researchers have experimented with less artificial methods to target children’s reasoning about kinship. Chambers and Tavuchis (1976) incorporated photographs into a Piagetian-style questionnaire that was itself built into a board game. LeVine and Price-Williams (1974) asked Hausa children about the kinship relations that hold within their actual families, while Ragnarsdóttir (1997, 1999) also used children’s own genealogies, rather than definitions, in her research with Danish and Icelandic children. The study presented here also eschews definitions; we test children’s comprehension and production of kinterms using factual questions paired with photos of children’s actual kin. (See §4 for methods and rationale.)

In addition to this innovative methodological approach, our study advances earlier research on children’s kinship concepts in three major ways. First, it focuses on an Australian classificatory kinship system (see 3.1 for discussion). Thus far, the only researcher to have investigated kinterm acquisition in an Australian language is Bavin (1991), who was principally concerned with children’s understanding of the mappings between the egocentric kinterms and sociocentric subsection terms (which do not exist in Murrinhpatha). Second, the paper explores the acquisition of a system of ‘universal’ kinship, that is, one in which terms are infinitely extendable. In English, the language which most kinship acquisition studies have investigated, kinterms only refer to a limited number of positions close to ego in a genealogical tree. How might learning differ when kinterms can make reference to extremely distant kin, who are
perhaps not related genealogically? (In Hausa, certain terms are infinitely extendable, but LeVine and Price-Williams (1974) explicitly avoided these terms in their acquisition study.) Third, this paper is the first to comprehensively report on children’s acquisition of kintax—the morphological encoding of kinship relations. The only prior investigation to explicitly target kintax was conducted by Forshaw (2014; 2016), as part of his PhD research into children’s acquisition of Murrinhpatha’s bipartite verbal morphology. Forshaw examined the production of sibling inflected verb forms by seven children aged between three and seven years old, as responses to animated photographic stimuli. The children between 5 and 7 produced sibling inflected verb forms inconsistently, unless prompted by a Murrinhpatha-speaking research assistant, in which case they subsequently replicated the appropriate responses. The younger children struggled to complete the task. Davidson (2018) also found inconsistent usage of sibling inflected pronouns at a similar age within a corpus of child peer interactions (see 3.2). On the basis of Forshaw and Davidson’s results, for all the studies we report on here, we elected to set the lower age bracket at 5 years, and to restrict our kintax study to comprehension only.

As summarised above, earlier research on children’s acquisition of kin terms has investigated the extent to which semantic complexity, as measured by number and type of compositional ‘features’, explains children’s performance on definition-based tasks. By contrast, in our project complexity is understood in terms of the conceptual structure of kinship systems; in particular, we are interested in how the notion of ‘distance’ affects children’s understanding of kinship relations. In other words, we consider the acquisition of different kin terms not in terms of their semantic complexity as individual lexical items, but in terms of their position in a relational system. We use two different measures of distance based on the anthropological literature on Australian kinship, as explained in 5.3.2.

Before turning to our Murrinhpatha acquisition studies, we provide some background about the Murrinhpatha language, its speakers, and its kinship system. In our discussion of Murrinhpatha kinship, we focus on those dimensions of the system which we target in our study, though we note that Murrinhpatha children must learn a great deal more about kinship than we investigate here.

3. The Murrinhpatha Language and Murrinhpatha Kinship.
Murrinhpatha is a headmarking polysynthetic language spoken by around 3000 people residing in Wadeye and surrounding communities, in the Moyle and Fitzmaurice Rivers’ region of Australia’s Northern Territory. Originally established in 1935 as the mission of Port Keats, like many of the Northern Territory’s remote communities, Wadeye has since then experienced enormous population growth (Taylor 2010), as large families are the norm. Murrinhpatha is spoken as a lingua franca by all Aboriginal people in this region. Through membership of a patrilineal clan (Falkenberg 1962; Ward 1983; Ivory 2009), Aboriginal people in this region claim affiliation to the Murrinhpatha language, or to one of the neighbouring languages: Marringarr, Marri Tjevin, Marri Ammu, Magati Ke, Jaminjung and Ngankiwumirri. While in the 1970s and 80s these neighbouring languages were well represented in Wadeye, they are now spoken by only a handful of elderly speakers, if at all. All Aboriginal people in Wadeye speak Murrinhpatha on a daily basis although certain lexical items and various structural features from these neighbouring languages have entered into modern Murrinhpatha usage (Mansfield 2015; 2017; 2019). The non-Aboriginal Australians (ku bamam, ‘white people’) of Wadeye comprise about 10% of the population (Australian Bureau of Statistics 2016). These mostly transient residents (teachers, nurses, police, etc.) don’t speak Murrinhpatha and are generally not considered kardu darrikardu (‘kin’/‘countrymen’). They are normally allocated kinterms only after displaying long term commitment to the community. Virtually all Aboriginal children in Wadeye enter school as monolingual speakers of Murrinhpatha, at five or six years of age, which is when they get their first sustained exposure to English. However, high levels of absenteeism and low school retention rates see many youth depart school with little command of English and low levels of literacy and numeracy (Taylor 2012).


Classificatory kinship systems can be thought of as systems that extend kin terminology far beyond close genetic kin and affines. Systems that extend indefinitely have been found in Asia, Africa, South America and Australia (Barnard 1978). All Australian kinship systems are bifurcate merging, in that relatives on the father’s side are distinguished from relatives on the mother’s side and relatives of the same sex (e.g. mother and mother’s sister) are categorically merged. The ‘unlimited’ nature of Australian classificatory kinship systems was first remarked upon by Radcliffe-Brown (1930/1931) who noted that the terminological equivalence of same-
sex siblings could potentially extend indefinitely. This feature, and the incorporation of affinal kin into the same kinship categories as consanguineal kin (step-kin merger, Scheffler 1978:103) give Australian systems their ‘universal’ character. Although challenged by proponents of categorical (monosemic) theories of kinship (Leach 1958; Needham 1974), structural mergers are key to extending non-focal referents from focal referents in genealogical (polysemic) accounts of kinterm semantics (Lounsbury 1964; Scheffler 1972; Scheffler & Lounsbury 1971). In the Australian context, merging principles are more than purely abstract equations, in that they have a practical application. Interlocutors compute their relatedness to hither-to unknown individuals by triangulating via a linking relative or connector (Dousset 2008) and then applying step-kin- or same-sex sibling mergers to the triangulated relationship (Blythe et al. 2019). This provides the requisite information for choosing appropriate reference forms and/or address terms, and for establishing expectations surrounding appropriate behavior. Same-sex sibling mergers also underpin sociocentric divisions like subsections and sections (the so-called skin systems). These systems streamline the process of calculating via linking relatives because persons within the same subsection are generally considered to be classificatory siblings. Although the Murrinhpatha do not have a skin system, same-sex sibling merger is nevertheless a key organizational principle in Murrinhpatha’s polysynthetic grammar and in its system of nominal kinship terms.

As we see in Table 1, the Murrinhpatha kinship system is organised along four patrilines. Referents of each kinterm are given in the corresponding row on the right side of the table. Because the Murrinhpatha system is structurally intermediate between the four patriline Aranda and the two patriline Kariera systems (Radcliffe-Brown 1930/1931), it was previously thought to be a system in flux (Stanner 1936; Falkenberg 1962). However, over the last 80 years the system has proven to be structurally stable (Keen 2013; Blythe 2018). Although the ‘system’ appears to be stable, various kinterms do show signs of possible lexical replacement. Some of these (mama, dedi, kas) are borrowings from English; others are innovations that derive from other sources (e.g. paba). ‘Junior’ terms are those especially used by children and by adults in child-directed speech, although certain adults use them all the time. Junior terms show simplified phonotactic structures, such as consonant cluster reduction. Because these borrowings and/or junior forms show the same denotations as the full/traditional forms, we treat them as variants. Within
conversational corpora kinterm tokens are very frequently produced, by both adults and children alike (Blythe et al. 2019).

<table>
<thead>
<tr>
<th>Generation</th>
<th>Kinterms</th>
<th>Junior terms/ recent variants</th>
<th>Patriline(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G+2</td>
<td><strong>kangkurl</strong></td>
<td><em>kakurl</em></td>
<td>FF, <em>FMH</em>, <em>FFZ</em>, <em>FMZH</em>, <em>FFB</em>, etc.</td>
</tr>
<tr>
<td></td>
<td><strong>kawu</strong></td>
<td></td>
<td><em>MM</em>, <em>MFW</em>, <em>MMB</em>, <em>MFZH</em>, <em>MMZ</em>, etc.</td>
</tr>
<tr>
<td></td>
<td><strong>thamuny</strong></td>
<td></td>
<td><em>MF</em>, <em>MMH</em>, <em>MFZ</em>, <em>MMB</em>, <em>MFZH</em>, <em>MMZ</em>, etc.</td>
</tr>
<tr>
<td></td>
<td><strong>mangka</strong></td>
<td><em>maka</em></td>
<td><em>FM</em>, <em>FMZ</em>, <em>FFW</em></td>
</tr>
<tr>
<td>G+1</td>
<td><strong>yile</strong></td>
<td><em>dedi</em></td>
<td><em>F</em>, <em>MH</em>, <em>FB</em>, <em>FFS</em>, etc.</td>
</tr>
<tr>
<td></td>
<td><strong>pipi</strong></td>
<td></td>
<td><em>MMBS</em>, <em>MFZS</em>, <em>SpMB</em></td>
</tr>
<tr>
<td></td>
<td><strong>kaka</strong></td>
<td></td>
<td><em>MB</em>, <em>FZH</em>, <em>MMS</em>, <em>MMZS</em>, <em>MFS</em>, <em>MFBS</em>, etc.</td>
</tr>
<tr>
<td></td>
<td><strong>kale</strong></td>
<td><em>mama</em></td>
<td><em>M</em>, <em>FW</em>, <em>MZ</em>, <em>MFD</em>, <em>MMD</em>, etc.</td>
</tr>
<tr>
<td></td>
<td><strong>nginarr</strong></td>
<td></td>
<td><em>SpM</em>, <em>SpFW</em>, <em>SpMB</em>, <em>SpMBD</em>, <em>SpMBS</em>, <em>MFZD</em>, <em>MFZS</em></td>
</tr>
<tr>
<td></td>
<td><strong>ngaguluk; arluk</strong></td>
<td></td>
<td><em>SpMB</em></td>
</tr>
<tr>
<td>Generation</td>
<td>Term</td>
<td>Sex/Role</td>
<td>Sex/Role Abbreviations</td>
</tr>
<tr>
<td>------------</td>
<td>------</td>
<td>----------</td>
<td>------------------------</td>
</tr>
<tr>
<td>G0</td>
<td>munak, mumak</td>
<td>Z, MD, MZD, FD, FBD, MMDD, FFSD, etc.</td>
<td>MMBSD, FD, FBD, MMDD, FFSD, etc.</td>
</tr>
<tr>
<td></td>
<td>ngathan, paba, pule</td>
<td>B, MS, MZS, FS, FBS, MMMD, FFSS, etc.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>pugarli, kas, kaski</td>
<td>MBC, FZC, FMDC, MFBC, FMZDC, MMSSC, etc.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>nangkun</td>
<td>MMBDS, ♀H, ♀HB, ♂ZH</td>
<td></td>
</tr>
<tr>
<td></td>
<td>purrima</td>
<td>MMBDD, ♀W, ♀WZ, ♂HZ</td>
<td></td>
</tr>
<tr>
<td>G-1</td>
<td>muluk</td>
<td>♂S, ♂BS, ♀BS</td>
<td>MBDS, FZDS</td>
</tr>
<tr>
<td></td>
<td>newuy</td>
<td>♂D, ♀BD, ♂BD</td>
<td>MBDD, FZDD</td>
</tr>
<tr>
<td></td>
<td>wakal</td>
<td>any G-1 male</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>wakal</td>
<td>any G-1 female</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G-2</td>
<td>nginarr</td>
<td>MBDC, FZDC, ♀DH</td>
<td></td>
</tr>
<tr>
<td></td>
<td>kirdeng</td>
<td>♀SD</td>
<td></td>
</tr>
<tr>
<td>G-2</td>
<td>kangkurl, kakurl</td>
<td>♂SS, ♂ZSS, ♀SD, ♀SSD</td>
<td></td>
</tr>
<tr>
<td></td>
<td>kawu</td>
<td>♀DS, ♂ZDS, ♀DD, ♂ZDD</td>
<td></td>
</tr>
</tbody>
</table>
TABLE 1. Regular (binary) kinterms in Murrinhpatha. All terms are classificatory excepting the descriptive terms ngaguluk (SpMB) and kirdeng (♀SD). Greyed out strings may be further contracted through same-sex- or half-sibling mergers (see 5.3). Italicised strings are step-kin.

<table>
<thead>
<tr>
<th>thamuny</th>
<th>♂DS, [male dorminant siblings]</th>
</tr>
</thead>
<tbody>
<tr>
<td>mangka</td>
<td>♂DD, ♂BDD</td>
</tr>
<tr>
<td>maka</td>
<td>♂ZSS, ♂ZSD</td>
</tr>
<tr>
<td></td>
<td>♂SS, ♂ZSS, ♂SD, ♂ZSD</td>
</tr>
</tbody>
</table>

Certain kinterms have special social salience, for children and adults alike. Nangkun (H) and purrima (W) are actual spouses and potential spouses, as well as actual and potential siblings-in-law. Preferred marriage is to a matrilateral second cousin (nangkun, ♂, or purrima, ♀), although cross cousin marriages are also acceptable. The category term nginarr denotes ‘poison cousins’ (G ±1 kin in MMB patriline) who are strongly avoided. G+1 males in this group (MMBS, MFZS, SpMB) and females (MMBD, MFZD, SpM, SpFW) can also be referred to as yile and pipi respectively, terms that are also used for regular uncles (MB) and aunts (FZ). Production data from the experiments in this study (see 5.5) suggest that children learn to refer to these ‘poison’ relatives as nginarr before learning that the more neutral terms pipi and yile can also be applied. Taboos also apply on pronouncing the personal names of opposite sex siblings (and to a lesser extent, opposite sex cross-cousins, pugarli). On the other hand, social bonds between same sex siblings and same sex pugarli are particularly close, forming the collective basis for Wadeye’s infamous metal mobs (Mansfield 2013).

3.2. KIN-BASED MORPHOSYNTAX (KINTAX).

At least sixteen Australian languages are known to mark kinship contrasts in their pronoun systems (Hale 1966; Schebeck 1973; Hercus & White 1973; Koch 1982; Alpher 1982; Evans 2003; Blythe 2013). Kin-based morphosyntax (or kintax, Evans 2003) generally reflects some categorical distinction, such as patrimoiety membership and/or generational harmony. These kinship contrasts may occur within free pronoun paradigms, and/or (in the case of the Murrinhpatha and Dalabon languages (Alpher 1982)) within paradigms of pronominal affixes to polysynthetic verbs. In Murrinhpatha ‘siblinghood’ is a morphological contrast, in dual and paucal number, encoded through the presence or absence of one of four fusional non-
sibling/gender/number markers. Thus within the duals, -ninha (DU.M.NSIB) and -ngintha (DU.F.NSIB) help distinguish two male non-brothers in 1 from two non-siblings that include at least one female in 2, respectively, and from a pair of siblings (brothers, sisters, or a brother and sister) in 3.8 ‘Sibling’ inflected verbs like parraneriwakthadharra lack these non-sibling/gender/number markers.

(1) danininthariwakthadharra (Blythe 2007:35)
    dani-ninha-riwak-dha-dharra
    3SG/DU.S.19Poke.PIMP-DU.M.NSIB-follow-PIMP-advancing
    ‘The two male non-brothers (♂ ♂) were following.’

(2) daninginthariwakthadharra (Ibid:35)
    dani-ngintha-riwak-dha-dharra
    3SG/DU.S.19Poke.PIMP-DU.F.NSIB-follow-PIMP-advancing
    ‘The two non-siblings at least one of whom was female (♀ ♂ or ♂ ♂) were following.’

(3) parraneriwakthadharra (Ibid:35)
    parrane-riwak-dha-dharra
    3DU/PC.S.19Poke.PIMP-follow-PIMP-advancing
    ‘The two siblings (♂ ♂, ♂ ♂ or ♂ ♂) were following.’

The presence or absence of the paucal non-sibling gender number markers -neme/-name/-nime (PC.M.NSIB) and -ngime (PC.F.NSIB) are implicated in the parallel contrast between ‘male non-brothers’, ‘feminine non-siblings’ and ‘siblings’, as illustrated by examples 4–6.10 The diachronic development of these siblinghood contrasts is covered in Blythe 2010 and 2013.

(4) parraneriwakthanamedharra (Ibid:35)
    parrane-riwak-dha-name-dharra
    3DU/PC.S.19Poke.PIMP-follow-PIMP-PC.M.NSIB-advancing
    ‘Several male non-siblings (♂♂♂♂♂♂) were following it.’
We illustrate how the expression of kinship distinctions often involves both the nominal kinterms and the verbal morphosyntax in examples 7 and 8, both of which contain the same reduplicated kinterm *kalekale*. Reduplicated kinterms denote close biological kin and/or plural referents. Example 7 comes from a story about a boating mishap in which some schoolchildren were nearly washed out to sea. In this case *kalekale* ‘mothers’ is coreferential with the subject of the verb *puddamkathukngime* (‘The mothers set off’). The verbal subject is 3PC.F.NSIB, which specifies the mothers as ‘more than two’ and as ‘not all being sisters’. These are the mothers of the schoolchildren. In contrast, 8 comes from a Djanba song that recounts events transpiring within the composer’s dream (Barwick 2011). *Kalekale* is coreferential with the dual sibling inflected verb *punnungamkangkatmirt=pumbankaya* (‘the two siblings were blocking [someone’s] way’). In this case the verbal morphology specifies the number of mothers as ‘two’ and the relationship between the mothers as ‘sisters’. In this case, the form *kalekalenukunu* is used for reference to a didgeridoo player’s mother and her older sister.
Davidson’s (2018) corpus-based studies of informal Murrinhpatha child-peer interactions reveal that unsolicited sibling inflected free pronouns are produced by certain children from as young as 4;6. However, inconsistency in their use of these forms suggests their comprehension of the siblinghood contrast at that age is not adult-like, although more systematic use of kintax begins to emerge about a year later. These corpus data show that the inception of kintax usage commences within interactions between family members including collectives of siblings and cousins. The experiment we describe in the next sections is designed to systematically test both children and adult’s comprehension of the kin-based verbal morphosyntax, including a conceptual distinction between parallel-cousins (who are classified as siblings) and cross-cousins (non-siblings).

4. THE DESIGN OF THE MURRINHPATHA KINSHIP ACQUISITION STUDY.

We set out to investigate Murrinhpatha children’s understanding of kinship language. We assessed the lexical and grammatical dimensions of children’s kinship cognition separately by means of the nominal kinterms tasks, described in §5, and the verbal kintax task, described in §6. As discussed in §2, previous research on children’s understanding of kinterms has largely relied on definitions. We avoided a definition-based approach, instead utilizing children’s own genealogies to determine at which ages, and how successfully, children can map genealogical relationships onto kinship terminology. Our aim here is to secure baseline data on kinterm comprehension and production, relative to age, widely sampling across the kinterm chart so as to vary close vs. more distant kin types. So doing, we aim to build up an age-graded picture of children’s kinship concepts, and see to what extent this picture aligns with the abstract principles of classificatory kinship identified by Radcliffe-Brown (1930/1931) and Scheffler (1978) (see discussion in §3).

Both the kinterms and the kintax tasks use photographs of living relatives within the subjects’ own genealogies. Mission records were used to build a genealogical database and photographs were taken of community members for use as picture stimuli. We estimate that our collection of 360 headshots represented 12.5% of Wadeye’s population (in 2015). Instructions
and audio stimuli were prerecorded (in Murrinhpatha). Audio stimuli, photographs and (in the case of the kintax task) animations were embedded within a Keynote presentation. The tasks were then administered using a tablet plugged into headphones and loudspeakers. All trials were recorded on video. All tasks were conducted with a close relative present who was requested to not answer questions on behalf of the child. Any responses by children that were prompted by co-present relatives were excluded from the dataset.

In §5, we discuss the kinterms experiment, and in §6, the kintax experiment.

5. The Kinterms Experiment.

The aim of this experiment was to test for children’s understanding of nominal kinship terms and their ability to map these terms onto real-life kin relationships.

5.1. Participants.

Twenty-four participants took part in the experiment (13 boys and 11 girls) ranging in age between 5 and 16. The mean age of the boys was 10, and of the girls was 11;2, while the mean age overall was 10;6. Participants were evenly distributed into four age bands, A-D, of 2 years and 9 months (see Table 2).

<table>
<thead>
<tr>
<th>Age band</th>
<th>Male</th>
<th>Female</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: 5 – 7;9</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>B: 7;10 –10;6</td>
<td>4</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>C: 10;7 –13;3</td>
<td>4</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>D: 13;4 – 16</td>
<td>2</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>11</td>
<td>24</td>
</tr>
</tbody>
</table>

Table 2. The demographics of participants undertaking the kinterms tasks.

5.2. Stimuli.

An individual list of referents and corresponding photographs were prepared for each child, including both consanguineal and affinal kin. Participants saw photos of individual referents presented on the tablet. Because all referents were long-term residents of Wadeye, we presumed
their faces would be recognizable to participants. As these customized tasks are time consuming
to build, certain participants conducted the same (or similar) tests as their biological siblings.
Each photo was a separate trial and no participant saw the same referent twice. The referent’s
genealogical distance from the child and other relationship metrics vary across the tasks. The
order of the stimuli for each child implies an increase in difficulty (see next section).
Our aim was to create stimulus lists for each individual participant in such a way that the
resulting list are comparable in terms of the range of contrasts tested and number of questions
testing each contrast. In practice, this is impossible, because all referent-participant relationships
are participant-specific and all families are different. As a result, there is some variation in the
types of kinship relations tested and the number of questions of each question type across
participants. We discuss this issue in greater detail in the Supplementary Information.

5.3. Design.  
Task Types.
Each participant undertakes three tasks. In the first, the Comprehension task, they have to
confirm whether or not the referent is related to them in the manner inquired about in the
stimulus question (e.g. *Kardu kanyika* [kinterm] *nyinyi?*, ‘Is this person your [kinterm]?’). 53%
of these questions are *apposite*, in that the proffered kinterm is appropriate for the relationship in
question. 47% of the questions are *inapposite*: the questions proffer kinterms known to be
inappropriate for the relationships in question.

In the second, the Production task, participants are asked how they would refer to the
referent (e.g. *Kardu kanyika ngarra namnathurran/namngethurran?*, ‘What do you say to/call
this male person/female person?’). In the third, the Altercentric Production task, they are asked
how the referent would describe them (e.g. *Kardu kanyika ngarra mambawurran?*, ‘What does
this person call you?’). Tasks 2 and 3 don’t explicitly request kinterms. In the absence of a
suitable vernacular label that could be used to elicit this class of words, the ordering of these
tasks after the Comprehension task primes the children for kinterm responses.
The Comprehension task begins with very simple questions (close biological relations: brother,
father, mother), alternating between apposite and inapposite questions. The task becomes more
difficult as it goes on, with genealogical distances increasing and as sibling and step-kin mergers
are introduced. The second and third tasks do not involve the same within-task shift in
difficulty. However, the three tasks are themselves ordered according to overall difficulty as predicted from the Piagetian studies (see §2), the third task being harder than the second task, and the second task being harder than the first task. The task design and format in the kinterms experiment is summarised in Table 3.

<table>
<thead>
<tr>
<th>Task</th>
<th>Task type</th>
<th>Question translations</th>
<th>Question type</th>
<th>Response types</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Comprehension</td>
<td>Is this your X?</td>
<td>Polar</td>
<td>Confirmations, disconfirmations, non-answers</td>
</tr>
<tr>
<td>2</td>
<td>Production</td>
<td>What do you call X?</td>
<td>Content (open ended)</td>
<td>Kinterms, names, descriptors, non-answers</td>
</tr>
<tr>
<td>3</td>
<td>Altercentric production</td>
<td>What does X call you?</td>
<td>Content (open ended)</td>
<td>Kinterms, names, descriptors, non-answers</td>
</tr>
</tbody>
</table>

**Table 3.** Tasks, question types and response types in the kinterms experiment

Tasks 2 (*Production*) and 3 (*Altercentric Production*) consist of information-seeking *content* questions. In contrast, the task 1 (*Comprehension*) questions are *polar* questions (see Table 3). While content questions are open-ended, polar questions (also known as *yes-no* questions) put forward a proposal about a certain state of affairs for the recipient to confirm as being either consistent or inconsistent with their own understanding (Bolinger 1978). The alternative term *yes-no questions* would suggest that responses to such proposals are restricted to *yes* or *no* answers. Under such an assumption, we might presume a *yes-no* ratio of 50:50 would apply in the context of a guess. However, responses that conform to the constraints of polar questions actually include confirmations, disconfirmations and non-answers. For task 1, attested confirmations include explicit ‘yes’ answers (*yu*), nods, repetitions of the proffered kinterm and combinations thereof. Disconfirmations included explicit ‘no’ answers (*awu, wurda*), headshakes, implicit disconfirmations (responses that instead propose alternative, non-equivalent kinterms), and combinations thereof. Non-answers included shrugs, *ya* (‘dunno’) and *mere tje mabatj* (‘I don’t know’).
A further reason why we cannot assume a yes-no ratio of 50:50 for guesses is that research within conversation analysis on polar questions reveals a preference for confirmations over disconfirmations (Sacks 1992; Pomerantz 1988; Heritage & Raymond 2012). What’s more, some Aboriginal people, especially children, have been found to gratuitously concur with polar questions within intercultural settings (especially in encounters with the legal profession), responding affirmatively when they don’t agree or don’t even understand the questions (Liberman 1980; 1981; Mildren 1999; Eades 1992; 2012; 2013). Consequently, we predict that respondents who are inclined to guess will say ‘yes’ – concurring with the alignment of the question. We thus expect the preferential bias for confirmations to favour concurring responses to apposite questions but disfavour concurring responses to inapposite questions, because blindly concurring with the question will bring about the wrong result.

In the comprehension task, responses were coded as correct if the participant accurately confirmed or disconfirmed the proffered kinterm. Responses were also coded as ‘non-answers’ if the participant responded with ‘I don’t know’ or a similar response (3% of responses). In the production and altercentric production tasks responses were coded as correct if the participant provided an appropriate kinterm. Unclear answers and non-answers were excluded (7% of all responses). The reported data have, per participant, 14-20 comprehension questions (half being apposite and half being inapposite), 6 production questions, and 3–6 altercentric production questions. Of 693 valid trials overall, 129 distinct individuals were referred to in the stimulus questions and depicted in the photographs.

**DISTANCE AND MERGERS.**

In the spirit of other kinterm acquisition research (e.g. Haviland & Clark 1974; Danziger 1993), we are interested in the extent to which children’s accuracy on the kinterms experiment relates to the complexity of different kinship concepts. To measure complexity, Haviland and Clark (1974) used a componential semantic analysis based on features such as [X PARENT of Y] and [MALE X] (see §2 for more details). In our study, we take a different approach and measure complexity on the basis of parameters identified within the anthropological literature on Australian kinship, namely, distance and mergers. We distinguish two possible ways of conceptualizing distance between kin, which we call GENEALOGICAL DISTANCE and KINSHIP DISTANCE. Taking an ‘etic’ perspective, we can calculate the conceptual distance between two individuals by counting the
relationships of descent that link one person to another, for example, a mother and a child have a distance of 1; a grandmother and a grandchild have a distance of 2; cousins have a distance of 3. This calculation is Genealogical distance: the raw number of nuclear kinship units that link ego to referent genealogically, prior to the application of merger rules.\(^{13}\) Kinship distance is a way of conceptualizing distance between kin that takes into account classificatory patterns of the Murrinhpatha language. Kinship distance refers to the outputted number of units of distance following the application of the merging rules. The mergers we consider when calculating kinship distance are same-sex-sibling merger, half-sibling merger and step-kin merger (Scheffler 1978).\(^{14}\) We use Genealogical distance, Kinship distance, and, in cases where the mergers make conflicting predictions about accuracy, each merger type as independent predictor variables for explaining our results.

Table 4 gives examples of how these factors were calculated. Tallies on the angle brackets in column 3 (>, >>, >>>) give the number of Same-Sex-Sibling, Half-Sibling and Step-Kin mergers, respectively (columns 4, 5 and 6). For example, in row nine of Table 4, the relationship FFBW (‘father’s father’s brother’s wife’ – a paternal grandfather’s sister-in-law) has a genealogical distance of 4 (four nuclear kinship units: F+F+B+W). However, FFBW can be reduced to FFW via a same-sex sibling merger: your father’s brother is terminologically equivalent to your father (yile). Then, FFW can be reduced to FM via a step-kin merger: Your father’s father’s wife is terminologically equivalent to your father’s mother (mangka). The mergers result in a kinship distance of 2 (two nuclear kinship units: F+M).

<table>
<thead>
<tr>
<th>Genealogical Relationship</th>
<th>Genealogical distance</th>
<th>Application of merging rules</th>
<th>Tallies on mergers</th>
<th>Kinship distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>M, F, B, Z, S, etc.</td>
<td>1</td>
<td></td>
<td>0 0 0 1</td>
<td></td>
</tr>
<tr>
<td>FF, FM, etc.</td>
<td>2</td>
<td></td>
<td>0 0 0 2</td>
<td></td>
</tr>
<tr>
<td>FMBD</td>
<td>4</td>
<td></td>
<td>0 0 0 4</td>
<td></td>
</tr>
<tr>
<td>MMBDS</td>
<td>5</td>
<td></td>
<td>0 0 0 5</td>
<td></td>
</tr>
<tr>
<td>MZ</td>
<td>2</td>
<td>MZ&gt;M</td>
<td>1 0 0 1</td>
<td></td>
</tr>
</tbody>
</table>
Recall that data on kinship distance, sibling merger, and step-kin merger are pooled from three different question types. For task 3 (altercentric production) questions, ‘egocentric’ responses are coded as ‘correct’ if the response displays an accurate understanding of the relationship in question (despite being an inappropriate answer to the question).15

5.4. Hypotheses.

The design of the experiment is based on two assumptions:

i. All referents can be categorized as the children’s kin, via principles of classificatory kinship.

ii. Children realize that potentially all (Aboriginal) persons within the community can be referred to with a kinterm.

We had the following hypotheses:

i. Recognising how genealogically distant relations may be classified as kin should be more difficult than with genealogically close relations. For example, it should be harder to identify your FMBS (genealogical distance = 4) than your FB (genealogical distance = 2).

ii. As factors determining a referent’s relatedness to ego, the abstract mergers (step-kin, same-sex- and half-sibling mergers) and their output (kinship distance), should affect the difficulty with which ego classifies the referent. Thus terminologically merged kin should be easier to identify than non-merged kin. For example, your FB and your F are terminologically equivalent (yile) while your FZ (pipi) is not equivalent to your M (kale).

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>FMFWD</td>
<td>5</td>
<td>FMFWD&gt;&gt;&gt;FMMD&gt;&gt;&gt;FM</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MZ&gt;FM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FFBS</td>
<td>4</td>
<td>FFBS&gt;FFS&gt;&gt;&gt;FB&gt;F</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>MH</td>
<td>2</td>
<td>MH&gt;&gt;&gt;F</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>FFBW</td>
<td>4</td>
<td>FFBW&gt;FFW&gt;&gt;&gt;FM</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>MFWD</td>
<td>4</td>
<td>MFWD &gt;&gt;&gt; MMD &gt;&gt; MZ &gt; M</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4. Exemplifying the coding of distance measures and merger rules.
23

So, it should be easier to identify your FB (kinship distance = 1, via same-sex-sibling merger: FB > F) than your FZ (kinship distance = 2). More mergers, however, should make kin identification harder. Thus despite your MFWD and your MZ both being terminologically equivalent to your mother (kale), the latter relationship should be easier to identify than the former because it implicates but a single merger (MZ > M), whereas the former relationship implicates three different types of merger (MFWD >>> MMD >> MZ > M) (compare rows five and ten of Table 4).

iii. The three tasks should get progressively harder. Selecting and producing the correct kinterm should be more difficult than merely recognising that a given kinterm may be applied to a referent. Identifying the appropriate term that referents will use to refer to ego will be the most difficult. It should be easier to determine whether a proffered term kaka (MB, etc.) may be appropriately applied to a specific male, than to choose kaka as the relevant term to apply to him when the term kaka hasn’t previously been mentioned in connection to that individual. Furthermore, it should be even harder to recognise that the person in question will probably not address you as kaka but will instead use the term wakal (‘nephew’, etc.)

iv. Older children should be better at all these tasks than younger children.

5.5, Statistical analysis.

We used Bayesian generalized linear multilevel regression to model the responses. We fit the model in R (R Core Team 2018) using brms (Bürkner 2017) and Stan (Stan Development Team 2018), specifying weakly informative priors (see the SI for details), a Bernoulli error distribution and a logit link function. The model estimates the probability of a response to a trial being correct or incorrect, depending on a number of independent predictor variables. The plots were created using ggplot (@ggplot2) (Wickham 2009).

Participant random intercepts are particularly relevant here since stimulus lists minimally vary across participants in the nominal experiment, as noted below. This means that the hierarchical structure of the data (answers grouped under participants) has a stronger effect on the results and mixed modelling allows us to take this explicitly into consideration. For more detail, see the Supplementary Information.
Each model was fit using a random intercept for participant and for referent image. These always improve model fit. The dataset is generally not large enough to fit random slopes. Model selection relied on the Watanabe-Akaike information criterion and 10-fold cross-validation. The models estimate non-linear terms of ordered factors; as a practice, we only report the linear term. Interactions were tested as appropriate; this is detailed for each case below.

The interpretation of ‘correct/incorrect’ is not consistent across the task types. The comprehension task has forced-choice questions, and a response is coded as ‘yes/no/non-answer’. The production and altercentric production tasks have open-ended questions, where a correct response is an appropriate kinterm. We expect a preponderance of affirmative responses in the comprehension task, irrespective of whether the question is apposite or inapposite (see discussion in 5.3). The apposite/inapposite distinction does not exist in the other two task types.

As a consequence, we report two separate models. Model 1 is fit on the comprehension task only and the response variable is whether the participant's answer is correct or incorrect. We explore the pattern of answers versus non-answers in the comprehension task separately. Model 2 is fit on the entire nominal experiment, and treats all responses as correct or incorrect irrespective of question type (forced-choice, as in comprehension, or open-ended, as in production).

MODEL 1 SETUP.
Model 1 is fit on 432 responses to comprehension questions from 24 participants to 100 distinct referents in the nominal comprehension task. We have the following predictor variables:

i. *Age band* (as an ordered factor). Participants were sampled in four age bands, from youngest to oldest: A (5-7;9) < B (7;10-10;6) < C (10;7-13;3) < D (13;4-16). Each band contains *six* participants.

ii. *Veracity* (as a binary factor): Apposite questions expect affirmative answers (confirmations). Inapposite questions expect negative answers (disconfirmations). There are *5-11* apposite and *6-11* inapposite questions per participant in the task.

iii. *Kinship distance* (as an ordered factor): the number of distance units between the referent and ego after step-kin or sibling mergers have been applied (see 5.3 above). Range: 1-5. The ranges vary across participants, with 3-7 questions per distance per participant.
The dataset is not large enough to consider genealogical distance and classificatory mergers separately. As a result, for this part of the experiment we only consider the classificatory distance units (kinship distance). Given the preference for affirmative answers, we tested for an interaction between age group and veracity.

**MODEL 2 SETUP.**

Model 2 is fit on the entire nominal experiment: 688 responses to questions about 134 distinct referents from 24 participants. We have the following predictor variables:

i. **Task type** (as an ordered factor): Comprehension < production < altercentric production. The tasks always follow in the same order and are considered as an ordered factor. There are 14–20 trials in the comprehension task per participant, 5–6 in the production task, and 3–6 in the altercentric production task.

ii. **Genealogical distance** (as an ordered factor): This is the descriptive distance between Ego and Referent, without considering any classificatory mergers. Range: 1–5. The ranges vary across participants, with 3–12 questions per distance per participant.

iii. **Step-kin merger** (as a binary factor): One of the classificatory mergers, discussed in Section 0. Varies across participants, with 23–28 questions without and 1–6 with step-kin merger.

iv. **Half-sibling merger** (as an ordered factor): One of the classificatory mergers, discussed in Section 0. Range: 0–3. Varies across participants, with 18–24 questions without, 6–7 with one, and 1–3 with two half-sibling mergers.

v. **Same-sex sibling merger** (as an ordered factor): One of the classificatory mergers, discussed in Section 0. Range: 0–3. Varies across participants, with 13–20 questions without, 10 with one, and 3–4 with two half-sibling mergers.

The task does not include kinship distance (which is the product of genealogical distance minus mergers) and question type (only applicable in the comprehension task) as predictors. We test for an interaction of age bands and task type: accuracy might improve more drastically with age in production versus comprehension, for instance.

**5.6. RESULTS.**
In line with our expectations, children’s responses in the kinterms experiment are graded by the child’s age, distance to referent, and task type. We first discuss the results of the comprehension task (model 1) and then of the entire task (model 2).

The estimates for the predictor effects in the comprehension task (model 1), along with 95% credible intervals, can be seen in Figure 1. This point-range plot shows the individual estimates for the intercept and the predictors, with 95% credible intervals. Put simply, the dot shows the single most likely estimate of the difference, while the line shows where 95% of the most likely differences lie. For the intercept, this is a difference from a 50% chance of a correct response. For the other terms, this is a difference from the intercept. If the entire line range excludes 0, we can be 95% certain that the difference is non-zero. For us, generally speaking, the non-intercept terms are of interest and we will focus on these throughout the analysis. When estimating the effects of ordered predictors, polynomial relationships between predictor and outcome are also considered. Given data scarcity, we only report and discuss linear (L) effects in the paper.

In Figure 1 three patterns emerge. First, the effect estimate for Age band is robustly positive, indicating a positive effect for age: older children are more likely to give correct answers. Second, the positive estimate for Question: apposite indicates that apposite questions are also more likely to have correct answers than inapposite questions, although the interaction of question type and age is not supported by model comparison; our model does not allow us to conclusively say that the ratio of correct responses to apposite versus inapposite questions improves with age. Third, the estimate for Kinship distance (the classificatory distance of Ego
from referent) is in the negative, indicating that a larger kinship distance makes a correct answer less likely. While the models consider correct, incorrect, and missing answers separately (ignoring the latter), our figures group correct and other answers together for the purposes of visualisation.

Figure 2 shows two-part representations of response accuracy per age band in the comprehension task. The left-hand panels show answers to apposite questions while the right-hand panels show answers to inapposite questions. The upper panels are scatterplots, in which each dot represents the mean accuracy for the given condition. Vertical line segments represent simple binomial standard errors. The lower panels are dodged bar plots. The height of each bar represents the absolute number of correct or incorrect answers in each category. The sum of the heights of the dark and light bars per category expresses the total number of answers in that category. The difference in their relative heights expresses the extent to which participants got the question right.

![Figure 2](image_url)

**Figure 2.** Participant age and question veracity in the comprehension task
What we see is that answers to apposite questions (the left-hand group) are overwhelmingly accurate, regardless of age, while answers to inapposite questions are more age-graded with responses becoming more accurate for older participants. The youngest group (A) is mostly guessing, and the oldest group (D) is overwhelmingly correct. We see a tendency that the age difference in response accuracy is driven mainly by the inapposite questions. This tendency is not robust (not supported by model comparison). These age trends are extrapolated in the overall kinterms experiment. Distance remains relevant, and the tasks clearly vary in difficulty. The overwhelmingly accurate responses to apposite questions confirm our prediction that younger respondents would be subject to the biasing effects of polar questions by concurring with the alignment of the questions, whether they know the answer or not. For this reason, evidence of conceptual development within the kinship domain is here demonstrated only in the age-graded improvement in the inapposite questions.

**Figure 3.** Model 2 estimates (nominal tasks). Linear estimates reported for ordered predictors.

Figure 3 shows the estimated effect sizes for the model of the entire kinterms task, with 95% credible intervals. The robust negative effect of Task type confirms our prediction that the production task is harder than the comprehension task, and that the altercentric production task is
harder than the production task. From the set of mergers, only *Same sex sibling merger* has a strong positive effect, which we discuss below in detail. *Genealogical distance* has a strong negative effect: referents that are further from Ego are harder to identify than close kin. Much like in the comprehension task, *Age band* has a strong positive effect across the board: older children are more accurate than younger children. We go through these effects in turn.

**Figure 4.** Age and task types in the nominal task.

Figure 4 shows results for the entire nominal task in a two-part representation. The upper panels show the mean accuracy of correct responses across age groups and task types (as scatterplots), while the lower panels show the total numbers of correct and incorrect responses (as dodged bar plots). The comprehension data are the same as in Figure 2, without the veracity-distinction. The total number of comprehension trials (left) is much higher than the number of trials in the other two tasks (mid and right), as shown by the height of the bars. In all three task types, children get more accurate with age. In the altercentric production task, only the oldest age group is able to provide a larger share of accurate answers.
Response accuracy is also related to Ego’s relationship to the referent. Greater raw genealogical distances between Ego and the referent result in lower response accuracy. In some cases, however, classificatory mergers have the effect of reducing these distances (as kin-distance), simplifying them in conceptual terms. One of these, same-sex sibling merger, has a robust effect of making children’s responses more accurate. We illustrate the process in Figure 5, which shows the effect of genealogical distance on response accuracy, and in Figure 6, where lineal kinterms are compared with collateral kinterms.

![Genealogical distance](image)

**Figure 5.** Accuracy across trials with genealogical distance.

The upper panel in Figure 5 shows mean accuracy across tasks for genealogical distance, while the lower panel shows the totals correct and incorrect. Accuracy gets worse with an increase in genealogical distance, but this does not convey the full picture. Another important dimension of these relationships is that classificatory mergers operate on them, which partly explains the sudden uptick in accuracy for genealogical distance of 5; most of the terms of this distance in the stimuli are subject to classificatory mergers.
FIGURE 6. On the left, accuracy across trials for all age groups for parents (F, M), parents’ same-sex siblings (MZ, FB), and parents’ opposite-sex siblings (MB, FZ). On the right, accuracy across trials for all age groups for siblings (B, Z), parents’ same-sex siblings’ children (parallel cousins; MZC, FBC), and parents’ opposite-sex siblings’ children, i.e. (cross-cousins; MBC, FZC).

The left-hand panels of Figure 6 shows accuracy for lineal kinterms in three categories: parents, parents’ same sex siblings, and parents’ opposite sex siblings. Accuracy is the highest for biological parents (yile, F and kale, M) who are easiest to identify (genealogical distance = 1, kinship distance = 1). Accuracy is comparable for same-sex parent’s siblings (FB, MZ) who are not distinguished from parents (genealogical distance = 2, kinship distance = 1, as a result of same-sex sibling mergers). By contrast, children’s accuracy is lower for parents’ opposite-sex siblings, for whom sibling mergers do not apply and who are distinguished from parents (pipi, FZ and kaka, MB; genealogical distance = 2, kinship distance = 2).

A similar pattern can be seen on the righthand panels of Figure 6 when we compare collateral terms across three categories; viz. cross-cousins (pugarli: FZS, FZD, MBS, MBD),...
biological siblings (ngathan, B and mumak, Z), and parallel-cousins (also ngathan: FBS, MZS; mumak: FBD, MZD). Recall that parallel-cousins are considered to be siblings, due to same-sex sibling merger. Biological siblings (genealogical distance = 1, kinship distance = 1) are relatively easy to identify, although less so than biological parents. Parallel-cousins (genealogical distance = 3, kinship distance = 1) are also easy to identify. Somewhat more difficult to identify are the cross-cousins (pugarli), for whom sibling mergers do not apply (genealogical distance = 3, kinship distance = 3).

5.7. DISCUSSION.
The scope of the experiment was necessarily constrained by the community demographics. Since referent images were tailored to individual participants, it was not possible to completely balance the design. Certain combinations of kinship distance or sibling merger simply did not exist within each participant’s genealogy. Because some of the participants are very young children, and because we wanted to keep the experiment similar across participants, the length and complexity of the experiment was restricted. However, these restrictions render our results very robust, since each child is tested with their own specific referents, and because we can compare children from a wide range of age groups. Our pioneering methodology has thus led to results that shed light on children’s ability to map Murrinhpatha kin terms onto real-world relationships.

In line with our expectations, the child’s age and distance from the referent, along with task type, determine success in the nominal task. Older children are more accurate in general, reflecting the gradual acquisition of Murrinhpatha kinship. While we have some evidence that the tasks are disproportionately harder for younger children, we do not have sufficient data to robustly confirm this finding. As for our distance measures, more distant referents are harder to identify in all tasks, but children are able to rely on same-sex sibling mergers to ‘reduce’ raw genealogical distance. This finding suggests that both genealogical distance and kinship distance are relevant to the cognitive processes involved in mapping kin terms onto kin; a model that takes into account both measures of distance better explains the results in the left panels of Figure 6, for instance. We discuss the cognitive implications of mergers in 7. With respect to task type, the comprehension task appears ‘easier’ than the production and altercentric production tasks, partly because understanding that a particular kinterm can be applied to a relationship is a precondition to actually producing the term in speech, and partly because
apposite polar questions (in which concurring guesses will be sufficient for success) present children with ‘free kicks’. Production questions, and particularly altercentric production questions are much harder, so much so that only the oldest children provide a large amount of correct answers in altercentric production. The evidence for the variable difficulty of the task type concurs with earlier findings on the role of perspective-taking in the acquisition of kin terms (see §2); the ability to take an altercentric perspective on kinship relations develops later than the egocentric perspective.

6. The Kintax Experiment.

We turn now to our second experiment. As outlined in §3, dual and paucal verbs inflect for siblinghood. The aim of the kintax experiment is to determine how well participants can relate grammatical forms to real-life kinship relations when they hear third person sibling or non-sibling inflected polysynthetic verbs.

6.1. Participants.

Thirty-nine participants took part in the kintax experiment: 23 were male and 16 were female (see Table 5). The youngest was 5 years of age while the eldest was 40. The 31 child participants were distributed into the same age bands, A–D, previously used in the kinterms experiments (2 years and 9 months apart). An additional ‘adult’ band (E) was added for the participants between 17 and 40 years of age. Of the children, the mean age of the boys was 8;11 and of the girls was 11;2. Overall the mean ages of the male and female participants were 11;9 and 16;5, respectively.

<table>
<thead>
<tr>
<th>Age band</th>
<th>Male</th>
<th>Female</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: 5–7;9</td>
<td>7</td>
<td>3</td>
<td>10</td>
</tr>
<tr>
<td>B: 7;10–10;6</td>
<td>6</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>C: 10;7–13;3</td>
<td>5</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>D: 13;4–16</td>
<td>1</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>E: 16;1–40</td>
<td>4</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>16</td>
<td>39</td>
</tr>
</tbody>
</table>
6.2. Stimuli.
In each trial, participants see one of five activities: fighting, laughing, pointing, walking, or waving. At the same time, they hear a sentence describing the activity with specific verbal morphology. Then they see two pictures, each containing photographic images of persons in the community. The configuration of persons in one of the pictures matches the verbal morphology of the sentence while the other does not. For example, a stick figure animation showing two women fighting is accompanied by a sentence with dual sibling morphology on the verb (Yawu piyemkathukparnamka, ‘Hey, two siblings are fighting’). One of the following images shows a pair of sisters; the other shows a pair of cross-cousins. Participants have to choose the correct picture.

Whether the people in the photos are siblings or not is available knowledge for our participants. This is because all 39 subjects taking the kintax experiment are descendants of two apical ancestors. Because they are all known to be blood relatives, and all are long term Wadeye residents, this allows a single standard test to be run for all subjects, using the same collection of photographs as stimuli. That is, unlike the nominal kinterms experiment, the kintax experiment was not tailored to each subject.

6.3. Design.
The experiment is conducted using a tablet with headphones for the participant and loudspeakers for the research assistant. Participants go through a brief training procedure and then answer a set of ten questions. There is then a short distractor game followed by another set of ten questions. Each participant undertakes twenty trials.

Murrinhpatha marks gender, number and siblinghood in its verbal morphology. The kintax experiment tests the ability to distinguish between siblings and non-siblings on the basis of a linguistic cue, and compares the uptake of siblinghood to the uptake of gender and number. Although we know adults use the siblinghood contrast to cross-reference biological and classificatory siblings (Blythe 2013), we know less about when children begin extending ‘sibling’ marking to classificatory siblings. As mentioned above, parallel cousins are classificatory siblings but cross-cousins are not classified as siblings. In our experiment we tested
three grammatical contrasts (gender, number and siblinghood), with ‘siblinghood’ involving two conditions. We test this cross- vs. parallel-cousin dimension of siblinghood independently from ‘other siblinghood’ (siblings vs. any other relationship). This gives us the following conditions:

Condition 1: GENDER all males vs. at least one female (3 trials per participant)
Condition 2: NUMBER dual (2) vs. paucal (3 or more) (3 trials)
Condition 3: X vs. // cross-cousins vs. parallel-cousins (7 trials)
Condition 4: OTHER SIBLINGHOOD siblings vs. non-siblings (7 trials).

So, in Condition 1, participants might hear ‘Several male non-siblings are laughing.’ Then they might have to choose between a picture of an all-male group and a picture of an all-female group (or a group comprised of females and males). In Condition 2, participants might hear ‘Two male non-siblings are fighting’, and then have to choose between a picture of two people and a picture of three people (or four people). In Condition 3, participants might hear ‘The two siblings are walking’, and then have to choose between a picture of two parallel-cousins (siblings in Murrinhpatha terms) and a picture of two cross-cousins (non-siblings in Murrinhpatha terms). In Condition 4, participants might hear ‘several male non-siblings are waving’, and then have to choose between a picture of a man with his father and his brother-in-law vs. a picture of three full brothers. In these illustrative examples, the first picture described is the correct choice, although in the dataset the ordering is randomised. An upshot of this design is that the other siblinghood questions (Condition 4) are relatively intuitive as many of them can be answered correctly with a general knowledge of Murrinhpatha society. For example, one picture might show a group of siblings while the other might show parents with their children, or a child with their grandparents – some combination that are clearly not all siblings. By contrast, both parallel-vs. cross-cousin pairs are of equivalent generations. This makes a ‘sibling’ reading at least plausible for each picture, meaning that more genealogical knowledge will be required to make the correct choice for Condition 3 questions than Condition 4 questions.

6.4. HYPOTHESES.

We had the following hypotheses about the results of this experiment:

i. Accuracy on the experiment will increase with age.
ii. The dual vs. paucal contrast will be harder than the gender contrast. We assume here that choosing a picture based on whether it shows males or females should be relatively straightforward, while choosing basis of counting (2 vs. 3 or more) should be a little harder (although perhaps not a lot harder). The most difficult morphological contrast should be siblinghood, because understanding siblinghood morphology requires more intimate social knowledge of the community – knowledge that is built up gradually through daily interactions.

iii. Participants’ ability to distinguish between parallel- and cross-cousins, and to identify parallel-cousins as siblings, should improve with age and experience. Older children will have a greater store of community level genealogical knowledge than younger children, and should better understand the categorical differences between parallel- and cross-cousins.

iv. Participants who are more accurate in the nominal kinterms comprehension task (see §5) should be more accurate in the kintax experiment. If a participant has developed a working understanding of kinterm semantics, and recognises how the individuals they encounter in the community are related to each other, they will be able to determine which people call each other ngathan (‘brother’), munak (‘sister’), and which people don’t, all of which should be requisite for accurate use of the siblinghood inflected morphosyntax.

6.5. Statistical analysis.

The modelling methods and the selection criteria were similar to those used in the nominal experiment. We fit two models: Model 3 was fit on all participants of the kintax experiment and tests for the effects of age and conceptual distinction on response accuracy. Model 4 was fit on participants who also took part in the nominal experiment and tests for the effect of accuracy in the nominal experiment on accuracy in the kintax experiment.

Model 3 setup.
Model 3 is fit on 780 responses from 39 participants in the kintax experiment. The model was fit with a participant random intercept. Referents do not repeat semi-randomly across participants the way they do in the nominal experiment so no referent random intercept was used. Given the
binary response structure of the trials, responses were coded as correct or incorrect. Otherwise, the same methods were used to analyse responses as in the kinterms experiment. Since we expected an interaction between age band and conceptual distinction, we tested this interaction. The model predicts whether the choice of a picture is correct or incorrect. The predictors are *participant age band* (as an ordered factor) and *morphological distinction*. We include this as a categorical variable. This is because our main focus is the relative difference of kinship-based morphosyntax as compared to gender- and number-based morphosyntax, and while there are reasons to assume a difference between gender and number as well, these are largely beyond the scope of this paper. We picked the gender contrast as the base level. The model estimates non-linear terms of ordered factors; here, we only report the linear terms.

**MODEL 4 SETUP.**
Model 4 was fit on 380 responses from 19 participants. The model was fit with a participant random intercept. The predictors were the same as in Model 3, with the addition of the participant’s average accuracy in the comprehension task of the nominal experiment.

**6.6, RESULTS.**
Much like in the kinterms experiment, participant age is an important predictor of success. The tasks also vary in difficulty. In Figure 7 we see the estimates of the model 3 fit on the kintax task.
FIGURE 7. Model 3 estimates (kintax task). Linear estimates reported for ordered predictors.

Much like in the nominal experiment, the effect estimate for Age band is robustly positive. Older participants are more accurate (in the kintax task, the oldest age group consists of adults and late teens). Responses to Number questions (Condition 2) and Other siblinghood questions (Condition 4) are very similar in accuracy to responses to Gender questions (our Intercept). Responses to Condition 3 questions (the cross- vs. parallel-cousin condition) are somewhat less accurate. The similarity between conditions 1-2-3 indicates that the morphological marking of siblinghood is not more difficult to learn than the marking of gender or number – this is not, strictly speaking, a grammar learning problem. Rather, the difference is higher-level: siblings and non-siblings were easier to discriminate than cross- and parallel-cousins. We go on to compare these two groups. We will take into consideration participant age as well, even though an interaction between age and distinction type is not supported by model comparison: that is, according to our model, all ages show the same shifts in accuracy across conditions. In Figure 8 we report the aggregated Siblinghood results, as based on 14 trials; that is, with conditions 3 and 4 neutralised.
Figure 8. Age and distinction type (referent gender, referent number, siblinghood) in the kintax task.

Figure 8 compares across participant age bands and morphological distinction types: referent gender (males vs. at least one female), referent number (dual vs. paucal), siblinghood (siblings vs. non-siblings). The upper panels show mean accuracy for each age band while the lower panels show the total numbers of correct and incorrect responses. For each distinction type, we see accuracy increasing with age. Across the board, the youngest participants perform at chance while the oldest participants are fairly accurate (c80% on average), regardless of distinction type. Evidently, children are not disadvantaged by the genealogical learning requisite for parsing siblinghood. The three morphological distinctions tested are apparently acquired at similar rates.

Figure 9. Participant ages against response accuracy in the kintax task for the separated dimensions of siblinghood (Conditions 3 and 4).

We further probe the different dimensions of siblinghood in Figure 9 by comparing participant responses across age bands for Condition 3 (the cross- vs. parallel-cousin condition) and
Condition 4 (Other siblinghood). As previously, the scatterplots display mean accuracy while the dodged bar plots display total responses correct and incorrect. The slightly lower rate of correct answers for Condition 3 than Condition 4 in age-bands A-D suggests children have greater difficulty discerning between groups of siblings and groups of cousins (who are of generational parity with siblings) than between groups of siblings and other types of non-siblings (and thus potentially of different generations). These differences however are slight and are not supported by model comparison: our model says that response accuracy changes the same way across condition types, across the age groups. For this reason, we feel the ‘aggregated’ siblinghood results (as depicted in Figure 8) give a more realistic representation of how this morphological contrast is used in Wadeye, where the sibling categories are understood as being classificatory (e.g. comprising both consanguineal and ‘merged’ relations).

For the 19 participants that took part in both the nominal experiment and the kintax task, age and accuracy in the kintax task are apparently strongly correlated (r = 0.6). Therefore, we refit model 3 on this subset (model 4a) and fit a separate model with participant average accuracy in the nominal comprehension task and morphological distinction (model 4b). Age band is, again, a robust predictor in (4a) (est = 0.86, est. error = 0.44), and the same is true for accuracy in the nominal task in (4b) (est = 2.43, est. error = 1.12, note that one of these is the linear estimate for an ordered predictor and the other is a numerical predictor, making these numbers difficult to compare directly). Neither model provides a better fit than the other, as indicated by model comparison using the Watanabe-Akaike information criterion and 10-fold cross-validation.

We conclude that accuracy in the nominal task is an important driving factor of accuracy in the kintax task, but that both are likely underlined by participant age.

6.7. DISCUSSION.

Regardless of distinction type, Murrinhpatha’s verbal morphosyntax is apparently difficult to master. Our task reveals gender, number and siblinghood contrasts being steadily acquired into adulthood. The additional genealogical knowledge required to accurately map real world relationships onto morphological distinctions does not result in developmental delays over and above that required for mastering gender and number contrasts. This genealogical knowledge is
seemingly acquired earlier, or in step with, the complex polysynthetic grammar. Genealogical knowledge is readily acquired in Wadeye, which is a small and close-knit community.

7. CATEGORICAL LEARNING OF LEXICON AND GRAMMAR.

Ours is the first quantitative experimental study of the acquisition of classificatory kinship principles and the first comprehensive study of the acquisition of kinship-related morphosyntax in any of the world’s languages. The kinterms tasks eschew definitions, instead tailoring individual sets of referents to each participant, enabling a direct assessment of individual children’s kinship knowledge. The kintax task is, to the best of our knowledge, unique in mapping children’s acquisition of kinship-related morphosyntax onto real world genealogical data, thus contributing to our understanding of how children learn to link linguistic forms with culturally specific meanings. Statistical power is constrained by our procedurally complex and time-consuming methodology and by the small population of Wadeye. We did not find robust statistical support for the interaction of age and task type in the nominal tasks, nor for age and distinction type in the kintax task. We owe this to the challenge of designing a balanced study that revolves around specific kinship relations and of expanding it to a large-enough sample of participants. Nevertheless, the tendencies remain relevant, providing useful insights into how children learn kinship-related lexicon and grammar in this community.

Across the various kinterm tasks, the performance of children in the youngest age band (5–7;9) show high levels of accuracy for close genealogical kin (93% for genealogical distance =1, 65%, genealogical distance =2), revealing that important kinterms can be mapped onto certain key individuals by five years of age. This result is supported by interactional data from the LAMP corpus that show children much younger than five spontaneously using relevant kinterms for their closest kin (Blythe et al. 2019). Taking into account all types of kin, children in the youngest age band (5–7;9) perform at chance levels in both the kintax comprehension task and in the inapposite kinterm comprehension questions. This suggests that although the focal elements of the classificatory system are in place by five years of age, the kinship categories underpinning the system are underdeveloped at this age. This assessment is supported by the inconsistencies in kintax usage reported by Davidson (2018) and Forshaw (2014; 2016). As our tasks are principally designed to track development relating to classificatory kinship concepts, the chosen five year lower limit seems well justified. On the other hand, the adult/late teen age band (16–40)
provides enlightening data on target performance in the kintax task. Unfortunately, the customised design of the kinterms’ tasks precluded introducing an adult/late teenage-band, as adults’ genealogies are not directly comparable with children’s genealogies.\textsuperscript{22}

In the nominal tasks, we see steady improvement in comprehension (in the inapposite questions) and production, and a late spike in altercentric production. These results, particularly that altercentric reference is harder to master than egocentric reference, are commensurate with the Piagetian investigations into perspective taking (e.g. Haviland & Clark 1974; LeVine & Price-Williams 1974; Greenfield & Childs 1977; Price-Williams et al. 1977; Bavin 1991).\textsuperscript{23} While perspective-taking abilities are clearly one important aspect of thinking about kinship, in our study we identified the conceptualisation of ‘distance’ to kin as a key learning problem for the classificatory system we’re investigating. Our approach adopts a genealogical (polysemic) analysis of kinterm semantics, in which lexical categories can refer to more than one genealogical position through the application of merging principles. As merging principles begin to consolidate, minimised kin-distances reduce the scale of the genealogical landscape, streamlining the mental calculus required for extending the small set of nominal kinterms (and the verbal kintax) out over the whole of society. Indeed the gradual accretion of merging principles seems to support a focal referent analysis of kinterm cognition, for at least most of the terms we examined from a developmental perspective. This is not to say we necessarily assume this to be the best account for all kinship systems, or even that it provides the best explanatory account for every term within the Murrinhpatha system. Conceptually, a term like \textit{nginarr} (see 3.1) seems more amenable to a categorical account (e.g. G ±1 kin in MMB patriline), in that it is hard to imagine specifically which subtype of \textit{nginarr} would occupy the focal referent position (see Table 1). If certain kin categories do emerge gradually by extension from focal referents, then perhaps the focal precursor is a salient individual in the community, rather than a prototypical position (Kronenfeld 2006) within a kin chart.

Another way of conceptualising ‘distance’, which we have not considered here, would involve measuring the sociocultural and/or experiential significance of a particular kin category. Benson and Anglin (1987) and Bavin (1991) suggest that child’s experience with kin can be a factor influencing development, while Ragnarsdóttir (1997, 1999) finds that demographic and cultural factors provide developmental advantages for Icelandic children over Danish children. Perhaps these experiential factors bestow similar advantages on children acquiring
Murrinhpatha. Wadeye, where all Aboriginal people speak Murrinhpatha, is also a small and homogeneous society. It could equally be argued Wadeye residents share an interest in genealogy and kinship, much like the Icelanders. Highly frequent social practices like the avoidance of opposite sex siblings’ personal names will provide children with salient genealogical input to feed their learning of kinship categories, build and refine their nominal kinship lexicon, and sharpen their use of kin-based grammar.24

Our data indicate that there is no lightbulb moment for children where kinship cognition or kintax miraculously falls into place. Both kinship relations and kin-based morphosyntax are learned gradually throughout childhood, into the teens and beyond. Our results also show that certain kinship relations are harder to master than others, and that siblinghood is not more difficult than gender and number. Yet for Murrinhpatha, this is perhaps to be expected. The distinctions tested here are all conveyed by the presence, or absence, of a portmanteau non-sibling/gender/number marker (see 3.2) – four morphemes to convey three sets of contrasts. And the distinctions themselves are not black and white. They are between lopsided categories of the type: X/non-X:

- **Siblinghood** = siblings vs. non-siblings
- **Gender** = males exclusively vs. females exclusively or females plus males
- **Number** = two vs. more than two

Thus none of these morphological contrasts is logically simpler than any other.

Our study poses some interesting questions for students of language and kinship alike. How much understanding of an abstract kinship system do children (or adults) actually require to adequately apply kinterms to the people they meet on a regular basis? And to what degree do children or adults need to understand the finer points of grammar in order to use it correctly, most of the time? Thus, if Murrinhpatha speakers are able to assume defaults, is it strictly necessary to know which people are or aren’t siblings? Of the 19 terms listed in Table 1, 17 are not sibling terms. Unless there is reason to think otherwise, ‘non-sibling’ would be a safe assumption.25 The higher cognition necessary for discerning more distal relationships in a ‘universal’ kinship system seems to consolidate long after fluency and grammatical competency develops.

How daunting would it be to learn genealogical relationships in a community of 3000 people, and how much knowledge is actually required of children? Our experiments do not
address the mechanisms of how genealogical knowledge is built up, nor how merging principles are specifically acquired. Nevertheless, there is reason to think that once merging principles kick in, rudimentary genealogical knowledge should be sufficient for adequate communication. A young boy going to play at a friend’s house might observe his older brother addressing this friend as kaka, and gleans that he too should address him as kaka, and his friend’s brothers as kaka; the same term he uses for his biological mother’s brothers. He also learns that his friend’s mother and father should be addressed with the grandparent terms kawu and thamuny (MBM and MBF, from his perspective). He doesn’t require specific genealogical knowledge about how his friend’s household is related to his own. Adults might possess such information but this advanced genealogy is irrelevant to him because same-sex sibling merger provides a shortcut (kinship distance) that is easier for him to process. Through play dates at friends’ (= relatives’) houses, categories like kaka expand beyond the set of immediate relatives, and it is these merging procedures that allow the expansion to proceed. This early understanding of kinship equivalences might explain why the siblinghood contrast is not mastered later than gender and number. Murrinhpatha morphosyntax is complicated and seemingly difficult to master, but this is not because it is kin-based. By the time a child can apply abstract grammatical knowledge to the psycholinguistic experiment we have set for them, rudimentary genealogies and basic merging principles are already water under the bridge.

That said, community level genealogical information is built up gradually with age and experience as social networks expand. Perhaps then it is unsurprising that we see continued improvement into the highest age bands for every dimension of our study. It would appear that learning to navigate within a universal system is probably a process of life-long learning. In this respect, acquiring ‘kinship’ in a place like Wadeye may have something in common with learning how to read and write, in that some of us continue to gain expertise throughout our lives, particularly with higher education and professional experience.

At this point we must acknowledge the clear limits of our methodologies, which were designed to determine what children at certain ages might be expected to know. A comprehensive understanding of children’s acquisition of the Murrinhpatha kinship system will require a multi-pronged approach which includes in-depth corpus-based work as well as ethnographic research into children’s daily lives—projects that are already underway. For our understanding of how children learn which terms to use when addressing their many kinspeople,
the LAMP corpus is proving an invaluable resource, revealing that at least some children receive quite explicit instruction about what to call genealogically close kin, and are corrected when they make errors (Blythe et al. 2019). Yet it remains something of a mystery whether the extended kinship categories are also learned procedurally, though similar processes of instruction. Though many questions remain, our study has provided baseline data on Murrinhpatha children’s understanding of kinship concepts and sets the stage for more qualitative approaches to grapple with the how of kinterm acquisition. Our study can inform a number of conversations in linguistics, anthropology and developmental psychology. We provide rare developmental data and analysis on a language from an understudied language family, broadening the focus from predominantly major world languages with millions of speakers to less well understood indigenous languages with fewer speakers, spoken in small societies where everyone knows each other. Our findings are also relevant to the study of kinship systems, especially to kinship acquisition, a field that has seen resurgent interest in the last couple of decades. They also contribute to acquisition research on understudied polysynthetic languages and of complex morphology, more broadly (Kelly et al. 2014). However, while Murrinhpatha may differ typologically from major world languages like English, and while the universal system underpinning its grammar and kinship lexicon is more expansive than kinship systems found in most urban industrialised societies, we would argue that our results point to the crucial importance of sociocultural knowledge on which the acquisition of all languages must rest.
APPENDIX

Tables of estimates for each model reported in the paper, with standard errors and 95% credible intervals.

Model 1

Correct response $\sim$ Age Band + Question Apposite + Kin distance + (1 | Participant) + (1 | Referent)

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Est.Error</th>
<th>Q2.5</th>
<th>Q97.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1.88</td>
<td>0.55</td>
<td>0.89</td>
<td>3.06</td>
</tr>
<tr>
<td>Age Band L</td>
<td>1.37</td>
<td>0.56</td>
<td>0.28</td>
<td>2.51</td>
</tr>
<tr>
<td>Age Band Q</td>
<td>-0.58</td>
<td>0.57</td>
<td>-1.78</td>
<td>0.49</td>
</tr>
<tr>
<td>Age Band C</td>
<td>0.22</td>
<td>0.55</td>
<td>-0.86</td>
<td>1.33</td>
</tr>
<tr>
<td>Apposite TRUE</td>
<td>1.05</td>
<td>0.52</td>
<td>0.05</td>
<td>2.09</td>
</tr>
<tr>
<td>Kin distance L</td>
<td>-1.32</td>
<td>0.71</td>
<td>-2.70</td>
<td>0.06</td>
</tr>
<tr>
<td>Kin distance Q</td>
<td>1.46</td>
<td>0.78</td>
<td>0.05</td>
<td>3.14</td>
</tr>
<tr>
<td>Kin distance C</td>
<td>0.60</td>
<td>0.55</td>
<td>-0.48</td>
<td>1.73</td>
</tr>
<tr>
<td>Kin distance E4</td>
<td>0.17</td>
<td>0.57</td>
<td>-0.92</td>
<td>1.27</td>
</tr>
</tbody>
</table>

Model 2

Correct response $\sim$ Age Band + Task Type + Genealogical Distance + Same-sex Sibling Merger + Step-kin Merger + Half-sibling Merger + (1 | Participant) + (1 | Referent)

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Est.Error</th>
<th>Q2.5</th>
<th>Q97.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.82</td>
<td>0.51</td>
<td>-0.15</td>
<td>1.81</td>
</tr>
<tr>
<td>Age Band L</td>
<td>1.38</td>
<td>0.60</td>
<td>0.22</td>
<td>2.57</td>
</tr>
<tr>
<td>Age Band Q</td>
<td>-0.13</td>
<td>0.58</td>
<td>-1.27</td>
<td>0.98</td>
</tr>
<tr>
<td>Age Band C</td>
<td>0.53</td>
<td>0.58</td>
<td>-0.61</td>
<td>1.65</td>
</tr>
<tr>
<td>Task2 L</td>
<td>-3.28</td>
<td>0.49</td>
<td>-4.31</td>
<td>-2.36</td>
</tr>
<tr>
<td>Task2 Q</td>
<td>-0.18</td>
<td>0.39</td>
<td>-0.97</td>
<td>0.57</td>
</tr>
<tr>
<td>Genealogical distance L</td>
<td>-2.21</td>
<td>0.71</td>
<td>-3.64</td>
<td>-0.83</td>
</tr>
<tr>
<td>Genealogical distance Q</td>
<td>0.29</td>
<td>0.52</td>
<td>-0.74</td>
<td>1.35</td>
</tr>
</tbody>
</table>
Genealogical distance C  -0.07  0.46  -0.96  0.81
Genealogical distance E4  0.27  0.46  -0.62  1.15
Same-sex sibling merger L  1.82  0.75  0.42  3.35
Same-sex sibling merger Q  -0.24  0.41  -1.06  0.59
Step-kin merger TRUE  -0.74  0.62  -1.97  0.43
Half-sibling merger L  -0.71  0.90  -2.56  1.00
Half-sibling merger Q  0.49  0.49  -0.45  1.46

Model 3

Correct response ~ Age Band + Distinction + (1 | Participant)

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Est.Error</th>
<th>Q2.5</th>
<th>Q97.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.97</td>
<td>0.23</td>
<td>0.52</td>
<td>1.41</td>
</tr>
<tr>
<td>Age bands L</td>
<td>1.16</td>
<td>0.24</td>
<td>0.70</td>
<td>1.63</td>
</tr>
<tr>
<td>Age bands Q</td>
<td>-0.01</td>
<td>0.25</td>
<td>-0.51</td>
<td>0.48</td>
</tr>
<tr>
<td>Age bands C</td>
<td>0.53</td>
<td>0.27</td>
<td>0.01</td>
<td>1.07</td>
</tr>
<tr>
<td>Age bands E4</td>
<td>-0.07</td>
<td>0.26</td>
<td>-0.59</td>
<td>0.42</td>
</tr>
<tr>
<td>Condition 2</td>
<td>0.10</td>
<td>0.30</td>
<td>-0.48</td>
<td>0.68</td>
</tr>
<tr>
<td>Condition 3</td>
<td>-0.46</td>
<td>0.25</td>
<td>-0.96</td>
<td>0.02</td>
</tr>
<tr>
<td>Condition 4</td>
<td>-0.10</td>
<td>0.26</td>
<td>-0.60</td>
<td>0.40</td>
</tr>
</tbody>
</table>
REFERENCES


BLYTHE, JOE.; ILANA MUSHIN.; LESLEY STIRLING.; and ROD GARDNER. 2019. Person reference and rights to know in Australian Aboriginal communities. Paper presented at 16th International Pragmatics Conference, Hong Kong.

BLYTHE, JOE.; JEREMIAH NGUBIDIRR TUNMUCK.; JEANNE MESSER.; and MARGARET CHI. 2019. Imparting the basics of Murrinpatha classificatory kinship. Presented at ALS2019, the annual conference of the Australian Linguistic Society, Sydney.


Haviland, Susan E.; and Eve V Clark. 1974. ‘This man’s father is my father’s son’: a study of the acquisition of English kin terms. *Journal of Child Language* 1.23–47.


From this point on we adopt the anthropological convention of referring to etic kin categories by initials: B = brother/brother’s, C = child, D = daughter/daughter’s, F = father/father’s, H = husband/husband’s, M = mother/mother’s, S = son/son’s, Sb = sibling, Sp = spouse, W = wife/wife’s, Z = sister/sister’s, ♂ = male/male’s [kin], ♀ = female/female’s [kin]. For example: ♂ZDC = male’s sister’s daughter’s child.

Although the classificatory kinship system appears complicated from an outsider’s perspective, and the process of acquiring it clearly takes time, we’re not suggesting that children immersed within these kinds of close-knit social networks are faced with exceedingly difficult learning problems. We are however suggesting that more expansive kinship systems are more difficult to learn. We thank an anonymous reviewer for this point of clarification.

Danish grandparent terms distinguish patrilineal from matrilineal descent: farmor, FM; farfar, FF; morfar, MF; mormor, MM; as do parent’s sibling terms: moster, MZ; morbror, MB; farster, FZ; farbror, FB.
Although sociocentric ‘skin’ systems are widespread within Australia (McConvell, Kelly, and Lacrampe 2018), many language groups, like the Murrinhpatha, have no skin systems. For this reason, we have chosen to investigate the merging principles that underpin all Australian classificatory kinship systems, as well as underpinning sociocentric skin systems like sections and subsections (see 3.1).


Trirelational kinterms (Blythe 2018) (or ternary kinterms, McGregor 1996) are not included in this table.

Murrinhpatha has a four-way number distinction: singular, dual, paucal and plural.

Morphological glosses used in this paper: DU = dual, F = feminine, FUT = future tense, M = masculine, NFUT = non-future tense, NSIB = non-sibling, PC = paucal (several), PIMP = past imperfective, PL = plural, S = Subject, SG = singular. 1, 2, 3 = first, second, third person.

Additional numbers between 1-38 convey verb class. For example, 3SG.S.19Poke.PIMP expresses the fusion of: third singular subject, 19 ‘poke’ verbal classifier, and past imperfective.

The past imperfective morpheme -dha is underlingly voiced. Murrinhpatha consonant cluster harmonisation is discussed in Mansfield (2019:103–109).

Whereas ‘masculine non-sibling’ denotes male only non-brothers, ‘feminine non-sibling’ denotes groups that include at least one female. ‘Feminine non-sibling’ is the unmarked (default) category, used for reference to families, households, etc., that are likely to comprise males and females.

The notation here represents years;months.

The original stimuli list contained two comprehension questions utilising trirelational kinterms (Blythe 2018). We excluded these from the analysis because these terms are more complex than the regular (binary) kinterms. Nevertheless, our suspicions were confirmed that even the eldest children would not recognise these terms. Garde (2013:119–120) states that the Bininj Gunwok kun-derbi trirelational terms are learned in adulthood or in the late teens. The Murrinhpatha trirelational kinterms are only used by the oldest generation of speakers. We thus predict they will soon fall out of usage altogether.

The ontological status of genealogical kinship trees has been the subject of much debate in anthropology (e.g. Schneider 1984). We use the genealogical model here as a heuristic device.

Scheffler’s same-sex sibling merging rule: ‘Let anyone’s sibling of the same sex as himself or herself […] be regarded as structurally equivalent to that person himself or herself; conversely, let any linking kinsman's sibling of the same sex as himself or herself be regarded as structurally equivalent to that linking kinsman himself or herself’ (Scheffler 1978: 115). The half-sibling-merging rule: ‘Let anyone's parent's child be regarded as structurally equivalent to that person's sibling (or parents’ child)’ (ibid: 102). The step-kin-merging rule: ‘Let anyone's parent's spouse (who is not also his or her parent) be regarded as structurally equivalent to that person's parent; conversely, let anyone's spouse's child (who is not also that persons’ own child) be regarded as structurally equivalent to that person's own child.’ (ibid: 103).

Let’s say a boy, seeing an image of his sister, responds to an altercentric question (‘What does this person call you?’) with mumak (Z). Despite the (egocentric) answer being incorrect, it still reflects an accurate matching of his relationship to the person. Thus, where relevant, accurate egocentric responses can be judged ‘correct’ insofar as they contribute to data on mergers. However, these responses are judged ‘incorrect’ when coding for altercentric production.
Anonymised raw data and code have been placed in the public repository [https://zenodo.org/record/3732882](https://zenodo.org/record/3732882).

Accuracy rankings between tasks are not directly comparable due to differences in task design, such as differing constraints on polar vs. content questions.

In the fighting, laughing, walking and waving stimuli, stick-figure animations convey movement. Stick-figure images were used in the ‘pointing’ stimuli.

The audio that accompanies the two images asks *Ngarranimin?*, ‘Which one?’. An ‘ear’ button on this slide allows participants, if necessary, to review the relevant animation stimulus and replay the audio stimulus.

For the first seven participants to take the task, one trial was eliminated as the audio quality of the stimulus was deemed inadequate. For these participants, 19 trials were included in the analysis. The audio was then rerecorded for all subsequent tasks.

We don’t have strong a priori reasons for assuming this particular hierarchical ordering. The ranking is a methodological utility in that logistical regressions require some variable to be set as the intercept. Departures from the intercept are used to detect differences from the hypothesized ordering.

Adults tend to have more G-1 kin than children, and fewer living G+2 kin than children. We did not use photographs of deceased relatives.

In particular, Bavin (1991:332) reports that by the age of 13, ‘only a few [Warlpiri children] could generalise beyond their own perspective.’ The eldest age band in Bavin’s study closely aligns with our band C (10;7-13;3). Our spike in altercentric production occurred in band D (13;4-16), which is older than the children in Bavin’s study.

One possible extension of the research presented here would be to develop quantitative measures that capture children’s everyday experience with kin and to test how well these measures predict children’s performance in the kinship tasks. Such an approach would need to be rooted in ethnographic research.

In fact ‘feminine non-sibling’ is safer still, as this category applies to families, most households and virtually any group that not exclusively male.