+ GEOGRAPHICA
HELVETICA +
supported by
sc|nat

Forests: the cross-linguistic perspective

Niclas Burenhult[1,2,7], **Clair Hill**[2,3,7], **Juliette Huber**[4], **Saskia van Putten**[5], **Konrad Rybka**[6], and **Lila San Roque**[5,7]

[1]Centre for Languages and Literature, Lund University, Lund, 22100, Sweden
[2]Humanities Lab, Lund University, Lund, 22100, Sweden
[3]Department of Linguistics, University of Sydney, Sydney, NSW 2006, Australia
[4]Department of General and Comparative Linguistics, University of Regensburg, Regensburg, 93053, Germany
[5]Centre for Language Studies, Radboud University, Nijmegen, 6525 HP, the Netherlands
[6]Berkeley Linguistics Department, University of California at Berkeley, Berkeley, CA 94720, USA
[7]Max Planck Institute for Psycholinguistics, Nijmegen, 6525 AH, the Netherlands

Correspondence: Niclas Burenhult (niclas.burenhult@ling.lu.se)

Abstract. Do all humans perceive, think, and talk about tree cover ("forests") in more or less the same way? International forestry programs frequently seem to operate on the assumption that they do. However, recent advances in the language sciences show that languages vary greatly as to how the landscape domain is lexicalized and grammaticalized. Different languages segment and label the large-scale environment and its features according to astonishingly different semantic principles, often in tandem with highly culture-specific practices and ideologies. Presumed basic concepts like *mountain*, *valley*, and *river* cannot in fact be straightforwardly translated across languages. In this paper we describe, compare, and evaluate some of the semantic diversity observed in relation to forests. We do so on the basis of first-hand linguistic field data from a global sample of indigenous categorization systems as they are manifested in the following languages: Avatime (Ghana), Duna (Papua New Guinea), Jahai (Malay Peninsula), Lokono (the Guianas), Makalero (East Timor), and Umpila/Kuuku Ya'u (Cape York Peninsula). We show that basic linguistic categories relating to tree cover vary considerably in their principles of semantic encoding across languages, and that *forest* is a challenging category from the point of view of intercultural translatability. This has consequences for current global policies and programs aimed at standardizing forest definitions and measurements. It calls for greater attention to categorial diversity in designing and implementing such agendas, and for receptiveness to and understanding of local indigenous classification systems in communicating those agendas on the ground.

1 Background

There is a well-documented discrepancy in how different official bodies define *forest* (Comber et al., 2005). Typically, however, similar sets of parameters are used for all such definitions, e.g. minimum physical requirements like "tree height" and "canopy cover" in combination with "land size". For example, the Food and Agriculture Organization of the United Nations' *Global Forest Resources Assessment 2010* (FAO, 2010) defines forest as "land spanning more than 0.5 hectares with trees higher than 5 meters and a canopy cover of more than 10 percent". In spite of this official definitional variability, there seems to be an assumption still that tree-covered environments are somehow a basic category that everyone can easily put a word to. Linguists too have long assumed that there is something fundamental about tree cover: the original Swadesh 200 Word List – a compilation of presumed basic, stable, and universal concepts for the purpose of comparison and historical reconstruction of languages – includes the item *woods* (Swadesh, 1952). Several subsequent comparative lexical studies do the same, such as Dyen et al. (1992) and Ringe et al. (2002). This reasoning presupposes that the concept expressed in English as *woods* or *forest* is a natural one for languages to delimit and instantiate lex-

ically, and that it ought to be translatable across languages with little effort.

However, the idea that some vocabulary is more "basic" or "universal" has long been under scrutiny (Hoijer, 1956). Furthermore, there is a growing appreciation within the field of lexical semantics that few concepts, if any, have identical lexical expression across all human languages (Majid, 2015; Evans, 2010; cf. Goddard, 2010), and that there are substantial challenges to cross-linguistic comparison of meaning (Haspelmath, 2010). In cross-linguistic studies, domains considered fundamental to human experience have received particular attention, such as motion (Talmy, 2000), the body (Enfield et al., 2006), and the senses (Majid and Levinson, 2011). These all point to considerable diversity in how meanings are lexicalized.

Landscape is another domain that has received increased semantic attention in recent years (Burenhult, 2008a; Mark et al., 2011; O'Meara, 2010; Rybka, 2016). This research shows that different languages segment and label the large-scale environment and its features according to astonishingly different semantic principles, often in tandem with highly culture-specific practices and ideologies. Presumed basic concepts like *mountain*, *valley*, and *river* cannot in fact be straightforwardly translated across languages. Yet, landscape features are a domain for which ideas of lexical basicness and universality persist, as in Chomsky's treatment of *river* as an elementary concept available to everyone independently of one's cultural and linguistic background (Chomsky, 2012) or in the computational semantic network analysis of Youn et al. (2016), which relies on traditional basic vocabulary, much of which is landscape-related.

Although an established topic of investigation in human geography and ecological anthropology, vegetation cover has so far taken a back seat in the linguistic inquiry into landscape. As with other landscape features, several questions come to mind. Which semantic principles and parameters do languages employ to classify tree cover? Is *forest* a universally recognized category? How do semantic strategies vary across languages? Existing literature does offer hints of an interesting diversity in how languages may lexicalize tree cover. For example, Lowland Chontal, a highly endangered indigenous language of Oaxaca in Mexico, is described by O'Connor and Kroefges (2008:299) as having a forest term *muña*, which apart from incorporating meanings similar to English "forest" or "jungle" also covers meanings like "bush", "underbrush", "overgrown wilderness" or "any type of weeds or garbage". Although *muña* is the closest Lowland Chontal equivalent of English *forest* or *woods* and a likely candidate for inclusion in a list of basic meanings, the term appears to invoke a wider meaning of a disorderly environment, one that is not kept in check by humans. An entity referred to as *muña* does not have to contain good-sized vegetation in the form of trees or in fact any vegetation at all.

An example of another semantic pattern comes from several indigenous languages in the Australia-Pacific region.

Here it is not uncommon to have a single term for "tree", "firewood", and "fire" (Schapper et al., 2016). For example, those are the English translations of the term *rowa* in Duna, a Papuan language spoken in New Guinea. The meaning of *rowa* also extends to instruments of fire, like "matches" and "lighter". Similar meanings are given for the term *thuma* in Wik-Ngathan and *yugu* in Guugu Yimithirr, two Pama–Nyungan languages of Cape York Peninsula, Australia (see Evans, 1992:495–496 for additional examples). The semantic connection between these concepts is obvious: fuel (firewood) and its primary material source (wood/trees) receive the same lexical expression as the burning of such fuel (fire). To our knowledge, no single language has so far been described as extending the meaning further to, say, "stand of trees" or "forest". However, cognate sets (i.e. sets of words with the same origin within a group of related languages) in some Australian languages provide evidence that such a semantic extension is not conceptually far-fetched. For example, in Kayardild – a Pama–Nyungan language spoken on the South Wellesley Islands, in the Gulf of Carpentaria – the word *thungkuwa* means "mangrove scrub" or "mangrove forest"; in closely related languages the cognates of *thungkuwa* have meanings like "wood", "firewood" (N. Evans, personal communication, cf. Evans, 1992). Indeed, familiar European languages sometimes display traces of co-lexicalization on parts of this semantic continuum, e.g. tree as land cover and tree as material, as in English *wood/woods* and French *bois*.

These examples underscore that the analyst must take maximal semantic diversity into account when approaching *forest*. The categories relevant to our inquiry may take us far beyond tree cover itself and its physical characteristics, and in very different directions according to the lexicalization principles of individual languages. As we explore further in the following section, it cannot be taken for granted that forest semantics can be straightforwardly defined, measured, or compared only in physical terms.

2 Tree cover: evidence from six speech communities

Our aim in this section is to identify and characterize semantically the closest lexical equivalent of *forest* in each of six languages, capturing some of the variation that may pose challenges to translations of the concept. The terms we investigate correspond to those that would be candidates for a list of basic vocabulary, such as the Swadesh 200 Word List (see Sect. 1). That is, they represent what we determine to be the best possible translational approximations of *forest*, and typically they are the translations offered by native speakers. We concede that this is a simplistic and selective approach to lexical semantics, and one that does not do justice to the complexities of language-specific lexical systems, nor their historical background. However, the approach does mimic and problematize the introduction of an external category and its

instantiation across human representational systems, thereby serving to highlight some of the pitfalls that *forest* is facing when confronted with linguistic and cognitive diversity.

In addition to exploring the meaning of forest terminology, we examine for each language the attendant question of whether treed environments are talked about as containers (that is, bounded entities which enclose other entities and which can be "entered" and "exited") or not. Containment is considered to be a fundamental topological relation which nevertheless displays cross-linguistic variation in terms of how it is semantically structured (Talmy, 1983; Bowerman, 1996). It has not yet been studied in depth in relation to potential landscape-scale "containers", such as forest. All languages in our sample give expression to this distinction in one way or another, offering interesting hints at cross-linguistic variation in how treed environments may be conceptualized.

Our sample is genealogically and geographically diverse, spanning six language families and four continents. Although the respective ecologies inhabited by the speech communities are also diverse, our sample does have a tropical bias. All six communities interact directly in one way or another with treed environments. All of them represent small-scale rural communities, typically follow traditional subsistence practices, and their lives are likely to be deeply affected by any forest-related policy. The accounts rest on long-term involvement in each speech community by the respective language expert and are based on a range of data types collected first-hand in the field, including stimulus-based tasks, elicitation, interviews, and natural occurrences of relevant terms in recordings. In each case, the analysis and interpretation involves a combination of such data types, resulting in the integrated accounts presented below. Methodological triangulation of this sort is recognized as best practice in language documentation and description (Lüpke, 2009).

2.1 A brief digression on the term *forest*

Before embarking on our global tour, we should briefly address the meaning of the term we have chosen to represent the starting point of our inquiry, namely English *forest*. The etymology of *forest* remains to be fully settled. It is likely that it derives via Old French *forest* from Late Latin *forestis* "royal wood, hunting ground protected by the king", but accounts of its earlier origins diverge. Gamillscheg (1969) proposes that the Latin term is derived from Frankish **forhist* (cf. German *Forst* "tended forest"), in turn derived from Proto-Germanic **furho* "fir", **furhísa* "fir forest" (the asterisks signify that the forms are reconstructed, not attested historically). An older analysis by Wartburg (1949) holds that the Latin form instead derives from *foras* "outside" (suggesting protected forest which is out of bounds for members of the public) or alternatively *forum* "court of law", implying that *forestis* evokes something subject to royal jurisdiction. Either

account shows that the term has gone through significant semantic change through history.

History aside, describing the present-day meaning of the term is far from uncomplicated. English is spoken as a first language by large numbers of people of diverse backgrounds and in vastly different environments, so speakers' notions of the meaning of *forest* are likely to vary across individuals and dialects. For example, an Australian English speaker's notion of *forest* may differ significantly from that of a British English speaker (cf. Bromhead, 2011).

For our present purposes we rely on several standard dictionaries in characterizing *forest*. What these all have in common is a definition of *forest*, which involves trees covering a large area. Some definitions are supplemented with references to high density and the presence of undergrowth. For example, the *English Oxford Living Dictionary* defines it as "a large area covered chiefly with trees and undergrowth", *Wiktionary* as "a dense collection of trees covering a relatively large area", and *Merriam-Webster* as "a dense growth of trees and underbrush covering a large tract". These definitions draw on parameters similar to the official ones (cf. the FAO, 2010 definition, Sect. 1). Trees and size of the area covered by them are clearly essential to the English meaning.

2.2 Avatime – *lịŋwàfụ̀*

The Avatime are about 15 000 subsistence farmers settled in eight villages in the southern Akwapim-Togo Ranges in the Volta Region of southeastern Ghana, western Africa. Their language belongs to the Kwa branch of the Niger–Congo language family. The Avatime territory is located in the Dahomey Gap, a section of the Guinean forest–savanna mosaic ecoregion which extends all the way to the Atlantic coast and interrupts the belt of tropical moist broadleaf forest that otherwise extends along the western African coastal region. The hills are covered by mostly secondary forest interspersed with farmland, most of which is shifted regularly but some of which is permanent; the area also features savanna-like environments with high grasses and few trees. The Avatime traditionally practise shifting cultivation, clearing new land with slash-and-burn techniques every few years. The main food crops are hill rice, cassava, yam, and corn, supplemented by a number of fruits and vegetables. The farmland is often located quite far from the villages, and the Avatime make daily trips on foot to tend their crops.

The Avatime category to be investigated here is *lịŋwàfụ̀*, the term used to refer to tree-covered areas. The history of this term is partly obscure, but it probably has three components: a prefix *lị-*, which specifies which class of nouns the category belongs to, and a root *wàfụ̀*, which can be analyzed as a combination of the verb *ŋwà* "to clear bush" (alternatively the noun *aŋwà* "a type of thorny vine") and a form *fụ̀*, the meaning of which is unknown. *lịŋwàfụ̀* is the only word in the language which refers to a tree-covered area and is thus seemingly the translational equivalent of English *forest*.

However, the presence of trees is not central to its meaning, as it can be used to refer to any undeveloped or uncultivated place where grasses, bushes, or trees grow. More appropriate English translations might therefore be "bush" or "wilderness". For the purpose of this study, eight Avatime speakers were asked in their own language "What is *lìŋwàfù*?". Of these, only five respondents mentioned trees, three of whom also mentioned grass and weeds. Two respondents explained that *lìŋwàfù* is a place where people do not live. Three people responded that *lìŋwàfù* is the place where people go to make fields for cultivation (*ɔ̀nyɔ*). Two respondents were happy to use the term *lìŋwàfù* for an area with few or no trees – one of them was asked whether it is possible to say that *òdzògbè* ("savannah, grassland") is *lìŋwàfù* and responded yes. However, a third respondent contrasted *lìŋwàfù* with *òdzògbè*, and said that there is no *lìŋwàfù* where there is *òdzògbè*. When asked what people do in the *lìŋwàfù*, the most common answer (from seven out of eight respondents) was "go to farm". The second most common was "hunting" (five respondents) and four respondents mentioned one or more of "clearing the bush", "cutting trees", and "setting fire", all activities done to make a plot of land suitable for cultivation. Fields for cultivation (*ɔ̀nyɔ*) are made in the *lìŋwàfù* but are themselves not *lìŋwàfù*. However, because fields are always inside the *lìŋwàfù*, it is possible to say you are in the *lìŋwàfù* when you are in your field (see below for further discussion of containment).

In reference to particularly dense forest with big trees where people do not farm or have not farmed for a long time, Avatime speakers will use the phrase *lìŋwàfù tsìtsì* "old forest". Such referents would presumably all qualify as English *forest* but, again, it is not a perfect semantic match, since the English term in this case probably extends to treed environments which are not considered to be *lìŋwàfù tsìtsì* to speakers of Avatime.

In discourse the term *lìŋwàfù* almost always co-occurs with a postposition *mὲ* "inside", e.g. in the compound *lìŋwàfù-nε mὲ bagaa* "bush animals", literally "animals of the bush's inside" (-*nε* is the definite article). This is a characteristic which it shares with most other terms referring to landscape features, and it suggests that referents of *lìŋwàfù* are perceived as having containing properties. You can "go into" and "go out of" a *lìŋwàfù*, and "be inside" it.

2.3 Duna – *hi*

The Duna (Yuna) language is spoken by approximately 25 000 people in Hela Province, in the highlands of Papua New Guinea, and is classified as a Trans–New Guinea language (Ross, 2005; Haley, 2002). The Duna homelands are in the central cordillera of New Guinea, and cover a broad altitude range from about 400 m (along the Strickland River) to over 3000 m in the south and northeast (Robinson, 1999). Forest in this area is broadly classified as Central Range montane rainforest, with dominant species likely to include

beech, laurels, and conifers (e.g. hoop pine), and plentiful orchids, ferns, and mosses (http://wwf.panda.org/). While forest cover and type vary depending on the area and altitude, most people live in close proximity to densely treed areas. The majority of Duna people are subsistence farmers (homes and gardens usually being established between 1200 and 1600 m in altitude), with sweet potato as the main garden cultivar. A diet of gardened food and domesticated animals (especially pigs) is mixed with seasonal semi-cultivated tree crops (such as pandanus), other foraged forest foods (e.g. mushrooms, ferns), and game. Thus, people spend time in the forest as part of subsistence activities, but the intensity of this varies greatly from person to person (e.g. depending on location, gender, and skill set).

The simplest way to translate *forest* in Duna is likely with the word *hi*. However, the Duna word evokes a more general notion of outdoor, undomesticated space, as well as (potentially) the presence of trees. As an environmental term, *hi* prototypically refers to wild forest areas, usually those that rise above the zone of settlement and agriculture. Such areas are spiritually charged and associated, for example, with powerful spirits and rites of passage that saw boys transformed to men in the testing ground of an untamed environment. *Hi* seems to contrast in at least some respects to the term *phou*, which refers to swamps and grasslands. However, beyond this contrast, it is not clear how the concept of *hi* enters into the more specific environmental classifications of the Duna. As described in detail by Haley (2002), these environmental classifications center on features such as altitude, drainage, water features, soil type, and characteristic species, for example, recognizing two main altitude regions, a lower one, *wapia rindi* "low ground" (which favors the growth of fruit pandanus), and a higher one, *karia seke rindi* "mountain base ground" (that favors the growth of nut pandanus). The latter is probably more typical of *hi*, but the overlap is not complete.

In addition to referring to forested places, *hi* is also used with the sense of "outside", in contrast to internal spaces. For example, when one is inside a house one can speak of going to *hi* to undertake an outdoor activity such as washing or to investigate a noise one has just heard outside. However, *hi* does not appear to be commonly used in relation to the outside of smaller objects or containers (cf. Jahai *hüp* in Sect. 2.4), although this needs to be tested further. The syllable *hi* also occurs in the verb *hiwa-* "to go/come out, to leave (a dwelling)", which is very likely a related word. *Hi* is furthermore a signifier of a lack of domestication. For example, *hi* can form part of the name of older or wilder varieties of cultivars, such as *hi da*, roughly translatable as "wild taro". Similarly, domesticating a pig can be described as making it such that the thoughts or consciousness of the pig are no longer *hi* but are directed towards humans.

There is little linguistic evidence that the Duna think of a forest or *hi* as a container to be entered. In the available data there are no examples of *hi* occurring in combination with

the spatial term *kou* "inside" (e.g. to mean something like "the inside of the forest"), nor of *hi* occurring with the verb *uku-* "go into", (as in "go into the forest"). However, *hi* is found in phrases with the case marker *-ra*, which is used to mark concealed and internal locations (such as the inside of a bag). Thus, *hi* can be associated with hidden places and is indeed still used by Duna people as a place of retreat, refuge, and concealment in times of strife and flight.

2.4   Jahai – *hüp*

The equatorial rainforests of the Malay Peninsula are home to minorities speaking Austroasiatic languages. One of these is Jahai, a community of about 1000 band-based foragers in the Malaysian states of Perak and Kelantan and adjacent parts of Thailand's Narathiwat province. The mountainous territory which they traditionally roam is covered in equatorial Dipterocarp rainforest characteristic of lower mountain regions. Much of the forest cover is still primary, but there is ongoing logging activity in the southern half of the territory. The area lies entirely within the rainforest ecosystem, although large waterways, occasional rocky outcrops, and open river flats sometimes leave gaps in the forest cover. Jahai habitation traditionally consists of temporary camps of huts or lean-tos, typically under tree cover but sometimes also in open areas. Although predominantly subsistence foragers, the Jahai occasionally make swiddens. Jahai society is deeply embedded in the forest environment, which provides the constant scene for livelihood, mobility, social and kin relations, and beliefs. The shaded environment underneath the canopy is considered cool and healthy, protecting people from the heat of the sun and from the monitoring senses of supernatural beings above.

So it may come as a surprise that the Jahai language does not have a term which corresponds straightforwardly to *forest*. The closest equivalent is the term *hüp*. This is the term which Jahai consultants will respond with if asked, for example, to translate the Malay term *hutan* "forest, jungle" into their own language (Malay is the unrelated majority language and lingua franca of Malaysia, spoken as a second language by all Jahai). In fact, they provide this translation with such ease that one will be forgiven for thinking that its semantics are unproblematic and that the case is closed. However, closer examination will show that *hüp* is a relational noun, that is, its meaning expresses a spatial relation rather than an entity. Its closest lexical equivalent in English is the nominal reading of *outside* but with the pronounced notion of an outside region or realm (and not a surface, as in an outside wall). It represents the ordinary way of expressing a spatial relationship of non-containment, such as location outside a container like a quiver or a hut. It is also used for reference to regions outside a perceived realm of human communication, activities, and social relations, for example, a location outside the zone of a conversation (Burenhult, 2008b). A third context is reference to regions and locations outside bodies of wa-

ter. In all of these readings, *hüp* is conceptually and lexically contrasted with *klɛŋ* "inside", an opposite and equally relational noun used for reference to the insides of containers, dwellings, camps, conversations, water, and so on.

Given the densely forested nature of the area, any region referred to as "outside" at a larger scale is highly likely to have tree cover. This may well contribute to the confusion about the translational equivalents of *hüp*. However, none of the readings explicated above requires tree cover in order to apply and, in fact, none of them encode any vegetation at all. Whether permanent or not, treelessness does not have an influence on meaning: a large expanse of barren rock is still *hüp*, as are treeless river flats, and deforested areas. Saying, for example, of a logged area that "the *hüp* is gone" is completely nonsensical in Jahai, as is "let's plant new *hüp*". Any approximation of the English concept of *forest* will have to involve paraphrasing by means of the terms *jhũʔ* "tree" and *haliʔ* "leaf, foliage". The latter term in particular is used frequently to express whether an area is forested or not, as in *sɔc ka=haliʔ* "the foliage is gone", in reference to a logged area.

Unlike some of the other languages discussed in this paper, Jahai does not offer any evidence that being among trees is perceived as being contained by them. The language does not have expressions like "go into" or "be in/inside" a treed environment. This is in agreement with the notion of *hüp* as "outside realm", and with the fact that the forest is tantamount to the Jahai ecosphere. It is in striking contrast with Malay, the neighboring majority language, which has expressions like *masuk hutan* "to enter the forest" and *dalam hutan* "inside the forest". It is worth noting in this context that any official forest policy conveyed to the Jahai will be communicated in Malay and thus conjures an understanding of *forest* which seems fundamentally at odds with the Jahai notion.

2.5   Lokono – *konoko*

Our next example comes from Lokono, a moribund language of the Arawakan language family, spoken in the pericoastal areas of the South American nations Suriname, Guyana, and French Guiana. The Lokono people are settled agriculturalists, who also practise hunting, fishing, and gathering. The areas inhabited by the Lokono are typified by dry plateaus dissected by numerous creeks and rivers. The depressions formed by these watercourses are dominated by rainforest, large parts of which are seasonally flooded. In northern Suriname, where the data were collected, four general vegetation types prevail – forest, wood, scrub, and grass – each of which is found across three different ecotopes – land, marsh, and swamp – resulting in a kaleidoscope of environments (van der Eyk, 1957:37). Primary forest on land is highly diversified and reaches up to 30 m. Marsh swamp forests are lower and less variegated. Woods are lower than the forest and make up dense walls of thin trees, overshadowed here and there by larger specimens. Scrubs and grasses form closed

and open savannas on both land and marsh. These areas, populated with grasses, sedges, and scrubs, are often surrounded by "savanna forest" – a transitory form between the savanna and the forest. The Lokono villages are located on the border between the rainforest and savanna, usually in the vicinity of creeks.

The Lokono divide their environment into two main categories: *konoko* and *karhow*. We will here be mainly concerned with *konoko*. This term refers to tall vegetation with a canopy that provides shade and obscures vision. It is associated with humidity and coolness. At its densest and darkest, the *konoko* is called *konoko êbera* "forest's giant", a phrase with the noun *êbera* "giant", or *konoko anakhubo* "forest's midpoint", with the spatial noun *anakhubo* "midpoint". *Konoko* is associated with a specific type of spirit called *konokokoya*, venerated and avoided through a system of cultural practices (de Goeje, 1943; Roth, 1915). Places such as *konoko êbera* and *konoko anakhubo* are described as far removed from the village and suggest a continuum from the primary forest, into which the Lokono venture only rarely, to the more "tame" forest around the village, where trees are felled to make swiddens.

As long as there is tree cover that provides shade, the area is called *konoko*, even if it is seasonally or permanently flooded. The noun *konoko* does not receive any modifiers of the type "swamp forest" or "marsh forest" when referring to such waterlogged areas. The presence of water is fully compatible with the term *konoko*. However, *konoko* is not felicitous when talking about open, flooded savannas without trees or with dispersed trees that do not form a continuous canopy (e.g. stands of moriche palm, *Mauritia flexuosa*). Such areas are called *onêbera* "rain's giant", a combination of the noun *oni* "rain" and *êbera* "giant". Similarly, on dry land, *konoko* is limited to vegetation that has an umbral canopy. The forests, woods, and shrub on land can all be tall and dense, but if there is no canopy providing shade *konoko* does not apply.

For canopyless environments, the term *karhow* is more appropriate. It applies to grass and scrub savanna and is associated with higher lying areas, heat, dryness, and lack of shade. However, *karhow* can also apply to areas containing trees, some relatively tall, if the trees are dispersed or not tall enough to provide continuous shade. This is the case with the aforementioned "savanna forest". Such areas are sometimes referred to in Lokono with the compound *konoko karhow* "forest savanna".

Lokono offers first-rate linguistic evidence that its speakers conceptualize treed environments as containers. When talking about motion into *konoko*, the Lokono use the verb *kodonon* "enter containment", restricted to contexts in which an entity moves into a contained space (e.g. into a house or from a big river into a small creek). Moreover, *konoko* frequently appears with nouns encoding containment such as *loko* "inside" or *koboroko* "inside (of multiple objects, e.g. trees)". By contrast, when describing motion into the

karhow, the opposite verb *fotikidin* "enter non-containment" is used. This verb applies when an entity moves out to an open space (e.g. out of a house or from a small creek into a big river).

2.6   Makalero – *alah*

Makalero is spoken by some 7000 people on the south coast of easternmost Timor, in the nation of East Timor. It belongs to the Timor–Alor–Pantar family, a small group of Papuan languages thought to originate in New Guinea (Ross, 2005). The Makalero territory extends from the coast to the mountainous interior of Timor, and the area is variably covered with tropical dry forest and moist deciduous forest. Most Makalero speakers are subsistence farmers who own different types of agricultural land: swiddens, orchards with fruit-bearing trees, small vegetable plots, and rice paddies. Main crops include rice, maize, vegetables, and coconut.

The Makalero word corresponding to the English *forest* most closely is *alah*. There appear to be three aspects central to the meaning of this term: firstly, it denotes areas covered in trees and other woody plants; as one Makalero speaker puts it: "The *alah* is a place where trees live". This is a prominent aspect of its meaning and is the first thing mentioned by Makalero speakers when asked to define the term. In this dimension, *alah* can be said to contrast with *fereh* "grassland" or with *muʔa namar* "land that has been cleared of trees" and *muʔa hare-hareʔ* "clean land, land with no vegetation". Secondly, the vegetation in *alah* grows wild, as opposed to that planted by humans. In this respect, it can be arranged on a continuum of human intervention along with *ama* "field, swidden" or *ado* "orchard", *dana* "abandoned field where grasses and trees are growing again", and *alah kirin* or just *kirin* "old forest". It thus becomes clear that *alah* functions on two levels: on the one hand, it is a generic term referring to any type of wooded area as explained above, and on the other hand, it contrasts with (*alah*) *kirin*. On this lower level, *alah* refers to a not very dense forest cover with relatively young trees which is exploited and tended (but not created) by humans. *Kirin* is distinguished by having tall trees and by being untended and located far away from people. Thus, in effect, *alah* has two meanings – "forest" in the generic sense and "tended forest" – and only the context can disambiguate the two. Lastly, *alah* contrasts with *muʔa hopan* "flat or treeless land". This not only shows again that tree cover is central to the definition of *alah* but also that it is associated in particular with mountainous terrain.

The term *kirin* "old forest" calls for some elaboration. It also translates as "old", "valuable", "heirloom", "taboo", and may be used to describe any old and prized object. Elsewhere in the landscape domain, *kirin* is used with *ira* "water" in the expression *ira kirin* "permanent spring". On its own, however, *kirin* can only be understood to mean "old forest". In Makalero culture, old forests and springs (and some other landscape features) are associated with potent spirits (called

muʔa ki-ouar "the land's master"), which must be treated with the proper respect. They require regular sacrifices and trespassing in their domain is prohibited; visiting such places is safe only for prominent members of the clan which owns the land.

In Makalero, descriptions of location usually require that the location and located entity are accompanied by what is referred to as a locative verb, such as *isiʔ* "to be at" or *mutuʔ* "to be inside" (Huber, 2017). However, with some nouns which refer to natural, large-scale entities which are prototypical reference objects for locations, such verbs are optional in certain syntactic contexts (not unlike Australian English *to go bush* vs. *to go into the bush*). This class of nouns includes place names as well as landscape terms such as *larin* "mountain", *meti* "sea", and *ama* "garden". *Alah* also belongs to this class, which shows that treed environments are conceptualized as natural locations at the landscape scale. Interestingly, in those contexts when such nouns do occur with the locative verbs, *alah* always co-occurs with *mutuʔ* "to be inside", unlike the other nouns, which co-occur with the semantically more general *isiʔ* "to be at". Thus, like Lokono (Sect. 2.5), Makalero provides compelling grammatical evidence that location in treed environments, here represented by *alah*, is conceived of as containment.

2.7   Umpila/Kuuku Ya'u – *maalatha* and *thungkuyu*

The Umpila/Kuuku Ya'u dialect group is a Paman language (Hale, 1964) spoken on the eastern coast of Cape York Peninsula, Australia. It is severely endangered, having only a small number of elderly fluent speakers. Cape York Peninsula, a large remote peninsula located in the far northeast of Australia, was traditionally home to 40 or so distinct languages of the Pama–Nyungan persuasion – thus, languages of this region group with the main phylic classification of Australian languages. Cape York Peninsula is notable today as an important wilderness area, the largest in Australia and of global environmental significance for the diversity of habitats and endemic species found there. The coastal and inland ecology and topography of the Umpila/Kuuku Ya'u people's region – located around Lloyd Bay – is particularly diverse. It consists of a patchwork of ecosystems: coastline with rocky outcrops, islands and extensive reef; mangrove everglades; large estuary systems; plains with salt pans; tropical savannah; rainforest, and so on. Notable for the discussion of forests is the presence of the Great Dividing Range which runs along the length of the northeastern coast of the Peninsula. The eastern face of the range, particularly in the northerly part of Umpila/Kuuku Ya'u territory, is covered in thick rainforest and vine forests. This extends coastwards along the edges of the waterways. The territory in the south and inland of the range features areas of dry sclerophyll forest and tropical savannah.

While attempting to find the closest equivalent of *forest* in Umpila/Kuuku Ya'u, we run into a different kind of challenge. The language distinguishes two major types of tree cover, expressed lexically as *maalatha* and *thungkuyu*. *Maalatha* is tree cover predominantly composed of one tree species, e.g. eucalyptus forest, sandalwood forest (*Santalaceae*), and so on. One speaker explains in the local English-based lingua franca: *only one kind (of tree), not different different one*. These tend to be dry forests, but speakers are not attuned to this aspect of *maalatha* environments, since the core semantic element is uniformity of tree type. Speakers closely associate this term with hillside environments. The following Umpila speaker's commentary on the meaning of *maalatha* is typical and makes this association explicit: *maalatha mean when you go inside hill in proper scrub now, ngampula waathaka kani maalathaku nha'atha ampuyku* ("we all will go to the forest and will enter to search for possum"). This example also suggests that speakers of Umpila/Kuuku Ya'u perceive *maalatha* as environments which you "go into" (lexicalized with the verb *nha'ana*) and are contained by.

By contrast, *thungkuyu* refers to rainforest, typified by a diversity of plant life including large fruit-bearing trees, strangler figs (*Ficus* spp.), palms and ferns, woody vines, and so on. When talking of *thungkuyu*, Umpila/Kuuku Ya'u speakers repeatedly highlight the impenetrability and darkness of this environment. They focus on the consequences this has for mobility and access: *you can't go through*; *like the dark in the night inside lo thungkuyu*. *Thungkuyu* are places where people move carefully in groups through fear of becoming disoriented or lost. Indeed, the term itself appears to be originally derived from the term *thungku*, which denotes dark hues.

Maalatha and *thungkuyu* both require tree cover that is extensive, rather dense, and contains trees of considerable size (as opposed to a category *iicha* "thickets", "shrubs"). Thus, in a sense, both appear to qualify as equivalents of *forest* and they are basically translatable into English as "monodominant forest" and "rainforest", respectively. However, neither assumes a generic meaning which covers both types of treed environments, and there is no additional term which covers both categories.

Both *maalatha* and *thungkuyu* are closely associated with the spatial term *pakay* "inside, under, down". *Pakay* and its opposing term *kani* "outside, top, over" are conventionally used to refer to certain environmental zones. *Pakay* is associated with inland densely vegetated environments, because its meaning concerns containment.

3   Discussion

A comparison of our translational equivalents of *forest* offers some interesting observations. First of all, none of the investigated terms can be conclusively analysed as evoking the same meaning as *forest*, although two come close: Lokono *konoko* and Makalero *alah*, which both require tree cover and

which both have a generic reading applying to any densely treed environment. The Umpila/Kuuku Ya'u terms *maalatha* and *thungkuyu* also require tree cover, but they represent a more fine-grained distinction than in English. Together they refer to entities which would probably all qualify as *forest*, as defined here, but the language lacks an overarching *forest*-type term which ignores the species characteristics of the treed environment. The Avatime candidate *lịŋwàfù* takes a step away from these meanings, because it does not require the presence of trees but refers to "bush" or "vegetated wilderness" in a more general sense. Duna *hi* and Jahai *hüp*, finally, abstract away from trees and vegetation entirely, both carrying a more general spatial meaning of "outdoors", "outside", or "outside realm". Our sample languages thus appear to spread out along a continuum, with highly concrete tree-encoding meanings at one end and more general, abstract spatial meanings at the other.[1]

This semantic continuum is congruent with whether the languages treat treed environments grammatically as containers or not. The languages with terms that are more concrete and encode vegetation – Umpila/Kuuku Ya'u, Lokono, Makalero, and Avatime – offer grammatical evidence that referents of the terms are conceived of as containers. By contrast, the Duna and Jahai terms can encode abstract spatial meanings and do not present such evidence but instead seem to conceptualize the referents as the opposites of insides, namely outsides. Following from this, we might hypothesize that Duna and Jahai do not construe treed environments as bounded entities, whereas the other languages do. This would be in line with the reading of the Duna and Jahai terms as denoting an unbounded, infinite "outside", while the human sphere inversely represents the bounded entity and container.

Comparing ecologies, we find that such an interpretation makes some sense. Duna and Jahai are arguably those languages in our sample with speakers who live in closest proximity to forested areas, spend the most time in them, and are most dependent on them for their livelihood. Both communities are surrounded by tree cover, which dominates the physical world that is accessible to their senses. Avatime, Lokono, Makalero, and Umpila Kuuku Ya'u speakers also interact rather closely with treed environments but do so on a more intermittent basis, approaching and penetrating them from afar for particular purposes. For these communities, tree cover is but one of several ecotopes and it does not dominate

their perceived landscape. The Lokono and Umpila/Kuuku Ya'u ecospheres are particularly heterogeneous in this respect. It is perhaps no coincidence that such communities conceptualize forests as bounded entities, give them distinct labels, and express them linguistically as containers which can be "entered" and "exited". It appears as if our sample communities divide the same domain into their own specific categories that are optimal for their cultural and environmental profiles.

It is also interesting to note how the communities diverge in terms of their ideology and beliefs surrounding "pristine" treed environments. For the Duna, the Lokono, and the Makalero such environments are spiritually charged; among the Makalero they are also highly respected and subject to restrictions. The Umpila/Kuuku Ya'u consider them dark, confusing, and fearful. Conversely, the Jahai consider them a safe and healthy home, whereas treeless spaces are uncomfortably hot and ritually dangerous.

We should also briefly note that our sample offers no evidence that suggests that similar subsistence modes result in similar concepts of forest. For example, those two communities in our sample which are traditionally hunter-gatherers – Jahai and Umpila/Kuuku Ya'u – represent the two extremes on our semantic continuum. They are also ideological opposites in their approach to rainforest environments.

4   Conclusions

It should be evident by now that *forest* does not lend itself to clear-cut translation across languages. Our small but diverse sample shows that terms that may be assumed to be close equivalents of *forest* in fact vary rather fundamentally as to their semantic encoding. While some do closely approximate the English meaning of a densely treed area of some size, others refer to untamed vegetation in a more general sense (akin to *bush*); yet others do not encode vegetation at all but instead evoke more abstract spatial meanings of *outdoors* or *outside*. Thus, crucially, there are languages for which the referents of the term do not need to contain that component which is so essential to both official and English definitions of "forest", namely trees. Effective translation of *forest* into such languages requires innovative paraphrasing, such as "outside area with trees". Conversely, our sample also demonstrates that there are languages with basic terms that partition treed environments into more fine-grained categories than *forest*, for example on the basis of species composition. This would be tantamount to English having, say, two completely different terms for coniferous and deciduous forest but no generic term for forest subsuming the two. Consequently, translation of "forest" into such languages requires more than one term. Previous literature indicates that the semantic diversity does not stop there: as noted in Sect. 1, languages may have *forest* terms which incorporate the meaning of *garbage*, and per-

---

[1]Although this proposed semantic continuum helps to arrange the meanings of terms in our particular sample, we do not wish to suggest that it is inevitably relevant to basic forest terminology in other languages. We acknowledge the likelihood of forest terms with meanings that invoke rather different semantic parameters which may not be amenable to plotting on the concrete-to-abstract continuum proposed here. Further research and detailed semantic analyses from more languages are needed to determine whether the continuum is more widely applicable or needs to be adapted, and whether there are terms to which it is not applicable.

haps even *fire*. We leave it to the reader to ponder translational solutions to such semantic phenomena.

In this paper we have explored some of the cross-linguistic diversity applying to forest terminology. This diversity is not only of academic interest – it may also have a bearing on the practical aspects of forest policies and programs which rely on intercultural communication for their implementation. In particular, it may have consequences for current efforts aimed at standardizing forest definitions and measurements. It also suggests that great attention needs to be paid to categorial variation in designing and implementing forest agendas, and that receptiveness to and understanding of local indigenous classification systems are crucial to successful communication of those agendas on the ground (cf. Wartmann and Purves, 2014). Most importantly of all, we hope to have shown that linguistic diversity, although sometimes an obstacle to comprehension, can be a rich source of information and inspiration for scientists and policymakers alike.

References

Bowerman, M.: Learning how to structure space for language: A crosslinguistic perspective, in: Language and space, edited by: Bloom, P., Peterson, M. A., Nadel, L., and Garrett, M. F., Cambridge, Mam MIT Press, 385–436, 1996.

Bromhead, H.: *The Bush* in Australian English, Aust. J. Linguist., 31, 445–471, 2011.

Burenhult, N. (Ed.): Landscape and language: geographical ontology in cross-linguistic perspective, Lang. Sci. (special issue), 30, 2008a.

Burenhult, N.: Spatial coordinate systems in demonstrative meaning, Linguistic Typology, 12, 99–142, 2008b.

Chomsky, N.: The science of language: interviews with James McGilvray, Cambridge, Cambridge University Press, 2012.

Comber, A., Fisher, P., and Wadsworth, R. A.: What is land cover?, Environ. Plann. B, 32, 199–209, 2005.

de Goeje, C. H.: Philosophy, initiation and myths of the Indians of Guiana and adjacent countries, Archives internationales d'ethnographie, p. 44, 1943.

Dyen, I., Kruskal, J. B., and Black, P.: An Indoeuropean Classification: A Lexicostatistical Experiment, T. Am. Philos. Soc., 82, 1–132, 1992.

Enfield, N. J., Majid, A., and Van Staden, M.: Cross-linguistic categorisation of the body: Introduction, Lang. Sci., 28, 137–147, 2006.

Evans, N.: Multiple semiotic systems, hyperpolysemy, and the reconstruction of semantic change in Australian languages, in: Diachrony within synchrony: language history and cognition, edited by: Kellermann, G. and Morrissey, M. D., Frankfurt, Peter Lang, 475–508, 1992.

Evans, N.: Semantic typology, in: The Oxford handbook of linguistic typology, edited by: Song, J. J., Oxford, Oxford University Press, 504–533, 2010.

Food and Agriculture Organization of the United Nations (FAO): Global forest resources assessment 2010: terms and definitions, Working Paper 144/E, Rome, Forest Resources Assessment Programme, 2010.

Gamillscheg, E.: Etymologisches Wörterbuch der französischen Sprache, 2. edition, Heidelberg, Carl Winter Universitätsverlag, 1969.

Goddard, C.: The Natural Semantic Metalanguage approach, in: The Oxford handbook of linguistic analysis, edited by: Heine, B. and Narrog, H., Oxford, Oxford University Press, 459–484, 2010.

Hale, K. L.: Classification of Northern Paman languages, Cape York Peninsula, Australia: a research report, Ocean. Linguist., 3, 248–265, 1964.

Haley, N.: Ipakana yakaiya – Mapping landscapes, mapping lives: Contemporary land politics among the Duna, PhD dissertation, Canberra, Australian National University, 2002.

Haspelmath, M.: Comparative concepts and descriptive categories in crosslinguistic studies, Language, 86, 663–687, 2010.

Hoijer, H.: Lexicostatistics: a critique, Language 32, 49–60, 1956.

Huber, J.: Natural locations and the distinction between 'what' and 'where' concepts: evidence from differential locative marking in Makalero, Linguistics, in press, 2017.

Lüpke, F.: Data-collection methods for field-based language documentation. Language Documentation and Description, 6, 53–100, 2009.

Majid, A.: Comparing lexicons cross-linguistically, in: The Oxford Handbook of the Word, edited by: Taylor, J. R., Oxford, Oxford University Press, 364–379, 2015.

Majid, A. and Levinson, S. C. (Eds.): The senses in language and culture, The Senses & Society (special issue), 6, 2011.

Mark, D. M., Turk, A., Burenhult, N., and Stea, D. (Eds.): Landscape in language: transdisciplinary perspectives, Amsterdam, Benjamins, 2011.

O'Connor, L. and Kroefges, P.: The land remembers: landscape terms and place names in Lowland Chontal of Oaxaca, Mexico, Lang. Sci., 30, 291–315, 2008.

O'Meara, C.: Seri landscape classification and spatial reference, PhD dissertation, Buffalo, State University of New York, 2010.

Ringe, D., Warnow, T., and Taylor, A.: Indo-European and computational cladistics, T. Philol. Soc., 100, 59–129, 2002.

Robinson, R.: Big wet, big dry: The role of extreme periodic environmental stress in the development of the Kopiago agricultural system, Southern Highlands Province, Papua New Guinea, Master's thesis, Canberra, Australian National University, 1999.

Ross, M.: Pronouns as a preliminary diagnostic for grouping Papuan languages, in: Papuan pasts: cultural, linguistic and biological histories of Papuan-speaking peoples, edited by: Pawley, A., Attenborough, R., Golson, J., and Hide, R., Canberra, Pacific Linguistics, 15–65, 2005.

Roth, W. E.: An inquiry into the animism and folk-lore of the Guiana Indians, Smithsonian Institution Bureau of American Ethnology, Washington, US Government Printing Office, 1915.

Rybka, K. A.: The linguistic encoding of landscape in Lokono, PhD Dissertation, Amsterdam, University of Amsterdam, 2016.

Schapper, A., San Roque, L., and Hendery, R.: Tree, firewood and fire in the languages of Sahul, in: The Lexical Typology of Semantic Shifts, edited by: Juvonen, P. and Koptjevskaja-Tamm, M., Berlin, De Gruyter Mouton, 355–422, 2016.

Swadesh, M.: Lexicostatistic dating of prehistoric ethnic contacts, P. Am. Philos. Soc., 96, 452–463, 1952.

Talmy, L.: How language structures space, in: Spatial orientation: Theory, research and application, edited by: Pick, H. and Acredolo, L., New York, Plenum, 1983.

Talmy, L.: Toward a cognitive semantics, Vol. II, Typology and process in concept structuring, Cambridge MA, The MIT Press, 2000.

van der Eyk, J. J.: Reconnaisance soil survey in northern Suriname, Wageningen, WAU, 1957.

Wartburg, W.: Französisches etymologisches Wörterbuch, Volume 3, Tübingen, J. C. B. Mohr (Paul Siebeck), 1949.

Wartmann, F. and Purves, R.: Why landscape terms matter for mapping: a comparison of ethnogeographic categories and scientific classification, GIScience 2014: Eighth International Conference on Geographic Information Science, 192–194, 2014.

Youn, H., Sutton, L., Smith, E., Moore, C., Wilkins, J. F., Maddieson, I., Croft, W., and Bhattacharya, T.: On the universal structure of human lexical semantics, P. Natl. Acad. Sci. USA, 113, 1766–1771, 2016.