Language complexity and interlinguistic difficulty

Eva Lindström

Department of Linguistics, Stockholm University

evali@ling.su.se

Abstract

This paper explores the related but distinct issues of linguistic complexity and difficulty, as from the viewpoint of an adult learner. Language complexity is seen as an objective property of a system, which could in principle be computed mathematically, while difficulty is grounded in the particular person who experiences the difficulty, involving factors such as the linguistic categories present and the nature of their marking in the learner's own language. This reasoning will be illustrated with one non-Austronesian language, Kuot, and its three Austronesian neighbours, Nalik, Notsi and Madak, of north-central New Ireland, Papua New Guinea.

1. Background

Kuot is spoken in the linguistically most diverse region on earth: the country of Papua New Guinea has 5 million inhabitants and around 800 languages. In the islands and on the coast of the main island of New Guinea there are over 100 languages belonging to the Oceanic branch of the Austronesian family, a relative new-comer to the region at about 3,500 years ago. Remaining languages are subsumed under the label “Papuan”, which, however, is a negatively defined category, implying only that these languages are non-Austronesian and spoken in this region. They dominate in the interior of New Guinea, and around 25 such languages are scattered across northern Island Melanesia. They are assumed, in the main, to derive from languages spoken by the first settlers of the region who arrived more than 40,000 years ago (and presumably in some subsequent waves; Spriggs, 1997).

Kuot (Lindström 2002) is the only Papuan language of New Ireland in the Bismarck Archipelago, indicated by shading in Map 1. It borders on three Oceanic languages, Nalik (Volker 1998), Notsi (Nochi, Lesu; Erickson & Erickson 1991), and Madak (Malom variety; Lee 1978, 1989, n.d.), each from a different subgroup of Oceanic.¹

@@ MAP 1 HERE @@

For the last approximately 100 years, there has been a growing lingua franca in the area, the English-lexified creole Tok Pisin (similar to what is sometimes called Melanesian Pidgin or Neo-Melanesian). Essentially every member of the Kuot community is fluent in it, and many children speak no other language. However, children in neighbouring Oceanic-speaking communities typically do speak the traditional languages of their communities, and in fact Kuot children who play with Notsi children will know Tok Pisin and Notsi, but not Kuot. Yet the neighbouring communities are well-nigh identically configured in terms of population size, culture, history, and so forth – if anything Notsi is the smallest of them all, spoken in only two villages.

My impressions of the language ecology of this area derive from a total of around two years of linguistic fieldwork in the period 1997–2004, spent mainly in the Kuot village Bimun on the south-east coast of New Ireland. In that community, Tok Pisin is used in village meetings and church services, and in many everyday conversations, even in households where both adults are full Kuot speakers. Generally, people aged 35 and older are full speakers, 20–25-year-olds are semispeakers, and children are not acquiring Kuot at all (although there is an exception in a closeby hamlet where Kuot is consistently spoken; however, the children switch to Tok Pisin as their main language of interaction as soon as they are old enough to go and play in Bimun). There are often non-Kuots present; some are New Guineans who

worked at the former plantation nearby and stayed on, others have married in from neighbouring language areas or further away, and some are temporary visitors. Politeness therefore often requires the use of Tok Pisin or another language shared by the persons present..

Speakers of Nalik, Notsi, and Madak languages unanimously declare that Kuot is inordinately difficult, while their own languages are “easy”. Currently, few people who move into the Kuot-speaking area ever acquire an active competence in Kuot, although most learn to understand it. Conversely, Kuot speakers often acquire quite a good competence in neighbouring languages if they are in contact with the speakers. This appears to have been the situation historically too, although more non-Kuots knew some Kuot then; today, cross-linguistic communication does not require bilingualism other than in Tok Pisin.

Another factor behind the notion that Kuot is difficult is most certainly that the other three languages considered are related to each other (and to all other languages on the island, except Kuot). Even though they belong in different subgroups of Oceanic, they share a relatively recent common ancestor, and much lexical and grammatical material will be familiar from one language to the other. Nalik and Notsi are closer to each other than either is to Madak, but the distance between any two of them is not greater than that between English and German. To these, Kuot may be said to be

like Finnish or Hungarian: of a very different origin and with a very different structure to its grammar and morphology, but culturally and historically very much integrated with them, and having a large body of lexical loans. There are of course salient differences between the New Ireland situation and the European case, such as the fact that the New Ireland languages are small in numbers of speakers, and spoken in even smaller settlements of swidden horticulturalists in societies with very little social stratification. In spite of indications that personal mobility was rather limited traditionally, it seems likely that a larger proportion of speakers of each language would have been in contact with speakers of other languages than in Europe. The assumption of extended contact across language boundaries is supported in oral history; in the fact that clan territories straddle language boundaries; in the fact that a fair amount of kin terminology is shared between unrelated languages; and in the fact that there is what Malcolm Ross (1994) has called a phonological alliance in this part of New Ireland, with several shared phonological rules and restrictions (Notsi participates only to a limited degree in this Sprachbund).

The Ethnologue (Gordon 2005) gives the population for each language but the information is from different years and different sources and therefore hard to compare: Kuot pop. 2400 (2002); Nalik 5138 (1990); Notsi 1836 (2000); Madak (the language, not the family; cf. note 1) 3000 (1985). Note, however, that language boundaries constitute census boundaries and that

these figures should be taken to indicate the number of residents within the census division defined by each language, rather than the actual number of speakers.²

It deserves pointing out that the Oceanic languages in question are by no means simple from a linguistic point of view. And, of course, the fact that the Oceanic languages in central New Ireland are more similar to each other than any of them is to Kuot does not in itself constitute a reason why it should be more difficult for, say, a Notsi speaker to acquire Kuot than for a Kuot speaker to acquire Notsi. Yet the situation is not symmetrical.

2. Complexity and difficulty

So how could we investigate the claim that Kuot is more “difficult” than its neighbours?

The concept of *complexity* that I am operating with is structural, viewing language, in this case grammar, as a system whose complexity could in principle be computed mathematically. Although linguists may disagree on how to formulate the metric, it is still the case that, for example, irregularity in paradigms is inherently more complex mathematically than regularity, in

the sense that more rules are required to describe it, or more machinery of some kind, regardless of the descriptive framework.

Thurston (1987) does not make an explicit distinction between system-based and user-based perspectives, but writes in the context of learnability, and he expresses the concept of complexity quite well in the following passage:

By way of terminological definition, the simplest language would have a perfect one-to-one correspondence between a unit of meaning and the form encoding it. It would have no stylistic or sociolectic variation. It would have the fewest possible number of grammatically marked obligatory distinctions and a small morpheme inventory from which the full lexical inventory could be predictably derived by a small set of regular derivational rules. Taking full advantage of analogy, each rule would be applicable wherever it made logical sense; and there would be one rule per function. The most complex language, on the other hand, would have all the things we use to torture students of introductory linguistics — allophony, allomorphy, unpronounceable consonant clusters, gratuitous morphophonemics, and unpredictable suppletion. It would have an enormous morpheme inventory with many near-synonyms differing in slight shades of meaning with implications of degrees of formality and socioeconomic status. The

most complex language would also have a lexicon which relied heavily on a large number of opaque idioms. (Thurston 1987: 41)

(Note, however, that in this paper we will concentrate on morpho-syntactics and touch only briefly upon vocabulary and social factors.)

Difficulty, on the other hand, is related to complexity but differs from it such that while complexity is a property of a system, difficulty is subjective. That is, something is difficult or not difficult for a particular person. Difficulty therefore depends on the individual we take as our starting point. If I am Swedish and learning Estonian, it is very difficult as the two languages are very different; if I am Finnish it is a whole lot easier as many words and structures are closely related, quite independently of the complexity of the systems involved.

In principle, categories that are present in one's own language, or other languages one knows well, should be easier to acquire in a new language, such as definiteness and number marking (disregarding for the moment the form of the marking). However, this is not always the case, as for instance with gender, the assignment of which differs quite a bit even between fairly closely related languages.

In this paper, our imaginary learner is an adult speaker of one of Kuot's Oceanic neighbour languages, who wishes to acquire Kuot in order to use it in communication.

The definition of difficulty here is dependent on the adult acquisition perspective, and other properties of language, such as aspects of language use, would presumably require a different analysis. For example, irregular paradigms could be beneficial in language input processing for individuals who already have a high competence in the language, as some cases of irregularity can most likely help avoid ambiguity, shorten search times for a match in the mental lexicon of a listener, etc., depending on your favourite model of language processing.

Other authors have made very similar distinctions to that which I make between complexity and difficulty. Dahl speaks of system complexity (2004: 42–43) which is independent of use (p39), and difficulty (pp39–40) which is anchored in an agent. Miestamo (2005; this volume) distinguishes absolute complexity as a property of a language, and relative complexity which is defined with regard to a language user. I agree with the distinctions made by these authors.

It may be noted that Kusters (e.g., this volume) differs somewhat, making the concept of complexity agent-related, although it should be noted that his

agent, also an adult L2 learner, is depersonalised by the assumption of no prior linguistic or cultural profile.

3. “Counting difficulty”

In this paper I have in mind a simple left-to-right procedure, imagining the path through a clause from the beginning to end, checking how many choices are encountered at each stage, roughly according to the following list:

- how many categories are morphologically marked (obligatorily/typically)?
- how large a paradigm is there per category?
- how much variation (esp. irregularity) is there in each paradigm?
- how many of the grammatical categories are present in the learner’s language?
- how similarly are they marked in the learner’s language?

The first three of these relate to complexity, the last two to difficulty.

Even such a simple left-to-right method is not unproblematic. For example, if an earlier choice limits the selection of possible forms at a later stage in the clause, how should we count the number of choices at the second stage?³ And what about categories that are not really compulsory but without which the construction is unidiomatic? We will return to such issues below.

Here, I will not perform a full count, but illustrate the sorts of structures and paradigms we would have to take into account if we were to attempt this kind of analysis, and discuss the sorts of problems that arise.

A few caveats are in order. I have very limited experience of the languages discussed except for Kuot, and have relied on other sources. There is a published grammar only for Nalik (Volker 1998). For Madak and Notsi there are unpublished typescripts which the authors, Robert Lee and Lee & Laurinda Erickson, respectively, have kindly let me use. I have not had access to text in Nalik.

A hazard in all types of typological comparison is of course the fact that the number and types of categories postulated are analysis-dependent, and some of the summaries to follow represent my own interpretation of the materials available to me.

Table 1 gives a short overview of some grammatical categories and the numbers of forms marking them in the four languages we are considering, plus the creole Tok Pisin. Parentheses for plural marking on nouns indicate that there is a form to mark non-singular but it is a separate word. The plus sign (+) shows that there are forms for the category but that counting is not applicable; the ‘3+’ for Kuot verb classes is because there are three intransitive classes paired with four transitive classes.

@@ TABLE 1 HERE @@

Some of these categories will be presented in more detail below. First however, we will have an example of a transitive clause with two full noun phrases from each language (Tok Pisin excluded). Noun phrases and verb phrases are given in brackets.

(1) Nalik

[*A mun finau*] [*di wut buak*] [*a vaal*]
 ART NSG thief 3NSG come break ART house
 ‘The thieves came (and) broke (into) the house.’

(2) Notsi

[*Yia-ka tamat*] [*a-choka*] [*kariin ko*]
 DEM-INDEF man 3SG.SUBJ-spear huge fish

‘One man speared a huge fish.’

(3) Madak

[*la-va-kin*] [*di-t-kis* *gugu*] [*leng-kaxi* *atdi*]

NM-PL-woman 3PL-CN-sit make NM.PL-basket their

‘The women are sitting making their baskets.’

(4) Kuot

[illegible]

3FOBJ-check-3MSUBJ.NFUT A. sun/watch(f)

3M.APOSS.3SG

‘Adam checked his watch.’

The Oceanic languages all have SVO word order, while Kuot has VSO.

Kuot further cross-references objects of transitive verbs, while the others do not.

We will now turn to examine some subsystems of the grammars of these languages in a little more detail.

3.1 The verb phrase

This section investigates the variability in marking within the verb phrase. Appendix I gives more detailed data for each language. The Oceanic languages are similar to one another in that all have a sort of bracket system, where the verb phrase is introduced by an obligatory subject marker, followed by five to six slots for optional markers before the verb stem (for Madak it is made explicit that only three markers at a time are possible). Preverbal markers include tense and aspect, mood, adverbial markers, and negation. After the verb stem more optional material can be added, more or less bound to the verb; primarily serialising verbs, and Nalik and Notsi have transitivity suffixes. The languages differ from each other in the number of slots, the number of categories in each slot and to some extent the order of items, and in the forms that mark each category.

Kuot does not have the bracket system, as the subject marker is more closely integrated with the stem, and further occurs in different positions relative to the stem depending on the verb class, and object markers also occur in different positions depending on verb class and grammatical person; see Appendix III. Other markers indicate categories very similar to those of the neighbouring Oceanic languages but with some peculiarities. First, a number of enclitics may attach to the first constituent of a phrase; second, the continuous marker *mən*, as well as full adverbs, may go in various positions in the verb phrase.

Kuot and Nalik differentiate two tenses, non-future and future, while for the Notsi and Madak it is not clear to what extent the use of the markers indicated for their rather more elaborate tense distinctions are strictly required or merely possible.

In terms of complexity, the main difference between the languages is that many more post-verbal slots are reported for Nalik and Notsi than Madak or Kuot. In terms of difficulty, the verb phrase structure of the Austronesian languages would transfer quite easily from one language to another, although forms and details would of course have to be learned. Kuot, on the other hand, does not have a very complex verb phrase structure but the variable position of some markers will be unfamiliar, and the many subject marking strategies together with the fact of object cross-referencing a definite challenge to a speaker of an Austronesian language.

3.2 Pronominal systems

All the languages under consideration have inclusive/exclusive forms of non-singular pronouns, at least some degree of expression of dual number and sometimes trial or paucal. All have some degree of pronominal marking on prepositions. Differences include the presence or absence of a separate set of subject markers parallel with the independent pronouns, and how

many grammatical numbers are used. Full paradigms are given in Appendix II.

Nalik has the smallest number of pronominal forms among the languages studied here. As is common among Oceanic languages, subject markers are different from independent pronouns; the latter are also used in object functions. The only pronominal forms that interact with prepositions are those for alienable possession. Although dual and paucal numbers exist, they are used only rarely, and are based on non-singular (sometimes alternatively singular) forms. Inalienable possessive constructions have all but disappeared from use.

Notsi pronominal forms make up a large system, with some consistency across categories but not full regularity. In Notsi, independent pronouns are used as subject markers, and may function as the only exponent of the subject argument of a verb. There are some special forms used with prepositions which are also used as independent object pronouns, overlapping to a large extent with the subject/independent forms. Special forms are used for the benefactive, mostly derived from other forms except in the first person, and there is also a set of related benefactive forms used when food is the gift. Dual and trial forms contain components of the numerals *lua* 'two' and *tuul* 'three'.

Madak has a fairly large set of pronominal forms, differentiating independent forms for subject and object, and having a separate set for subject marking on the verb. Again, special forms go with particular prepositions, most of which are derived from other forms in the system. There are three numbers: singular, dual and plural.

The Kuot pronominal system has many forms, due in part to the fact that verbs cross-reference both subjects and objects. The forms used differ between the verb classes, and a class of adjectives take a partly different set of markers again (cf. also Appendix III). Prepositions are indexed with pronominal forms if there is no nominal in the prepositional phrase (and sometimes even if there is). Markers for inalienable possession are closely related to object prefixes, which are also nearly identical with preposition indexers. There is a large system of forms marking alienable possession, indexing all 12 pronominal categories of possessors, as well as possesseees (with some syncretism of masculine and feminine possesseees); forms for singular feminine possesseees are similar to subject forms of verb class I. There are three numbers: singular dual and plural, and two genders: feminine and masculine, distinguished only in the third person singular.

It is hard to say which is the most complex of these pronominal systems. Nalik and Madak have relatively small systems; the Notsi system makes more distinctions and has four numbers but recycles many of the forms

adding suffixes to form new categories, and does not distinguish subject forms from independent forms. Kuot expresses a large number of categories but recycles a fair amount of material between them, such that it is often possible to recognize the person and number of a form new to the listener. In terms of difficulty, Kuot causes speakers of neighbouring languages a lot of trouble by distinguishing feminine and masculine in the third person singular, by marking objects on verbs, and by using different positions of marking and somewhat different markers in different verb classes. Of these complications, it appears that gender is the hardest to master; even very fluent learners who handle verb classes and object agreement without thinking make frequent gender mistakes.

3.3 Nominal morphology and noun phrases

Notsi and Nalik have no nominal morphology. Number marking is done with a free morpheme.

Madak, on the other hand, has quite a large set of prefixes for nouns. They are prefixed articles (or noun markers), which can be combined with morphemes that mark plural and a number of other categories such as diminutive and parthood. The assignment of forms to particular nouns is not fully transparent, and noun classes (declensions) can be postulated among

common nouns on the basis of prefix selection; if based on the plural prefix there is one big class and four smaller ones.

Kuot nouns fall into 13 declensions based on plural formation. For about half the nouns, the plural is formed by suffixing *-(V)p* to the noun stem. Most of the rest end in particular forms which are removed before the plural ending is added (e.g., *kiraima* (sg) – *kiraip* (pl) ‘nail, claw’; *kikinəm* (sg) – *kikip* (pl) ‘ear’). The dual *-ien* is added onto the plural (except in many dyadic kin terms which have special dual forms).

Unpredictability and irregularity as in Madak and Kuot automatically give complexity as set out above. For speakers of Nalik and Notsi marking plural on nouns at all is an unfamiliar thing to do and they would have to acquire it as a linguistic habit so to speak. Speakers of Kuot and Madak expect plural marking but the irregularity would still cause problems.

Aside from morphology, the syntax of noun phrases appears to be roughly equivalent in complexity across all four languages.

4. Discussion

The above inspection of three grammatical subsystems in four central New Ireland languages points to one of the problems with this type of exercise: the different areas do not necessarily lead to the same ranking of languages. For instance, Kuot and Madak verb phrases are not very complex while nominal morphology is, and the opposite holds for Nalik and Notsi.

Further, it is not always clear which criteria should be emphasised; for instance the number of functions encoded in the pronoun system vs. consistency of form across each person-number category.

The left-to-right procedure in this crude form also fails to capture interdependencies between items in a clause, such as agreement marking on demonstratives and other constituents. Nor is it clear how to treat the variable affixation order that comes with the different verb classes in Kuot.

Regardless of how we determine complexity, however, it is clearly the case that much more than complexity figures into difficulty, and we will have one more demonstration. Let us imagine a Notsi speaker who is going to approach the production of a Nalik clause. Let us say he wanted to produce the clause given in (1) above:⁴

(1) Nalik

[*A mun finau*] [*di wut buak*] [*a vaal*]

ART NSG thief 3NSG come break ART house

‘The thieves came (and) broke (into) the house.’

We will assume that the lexemes are known. The word order is SVO as in Notsi so that is not a problem. Almost all Nalik nouns take a non-specific article (noun marker) *a* (or *na*) in most contexts, while Notsi has no articles, so our speaker may or may not remember to put that in. In the first noun phrase he would probably not have a problem with the plural marker *mun*, as Notsi has *no* in the same place and function.

In the verb phrase, Nalik subject markers are different from the independent pronouns, which may cause trouble. Now this is the sort of area that requires careful thinking if we wanted to calculate difficulty. Do we imagine our Notsi speaker confusing subject markers only with independent subject pronouns, or with third person plural forms from across the whole pronominal paradigm? As for serial verbs, there are similar constructions in Notsi, so again no problem. The last noun phrase has only an article and a noun so only the article would make the construction different from a Notsi phrase.

Now let us imagine our Notsi speaker attempting a Kuot clause, the one given above as (4):

(4) Kuot

[*o-ikat-on*] [Adam] [*muabari an*]

3FOBJ-check-3MSUBJ.NFUT A. sun/watch(f)

3M.APOSS.3SG

‘Adam checked his watch.’

Here the word order is VSO, which will be unfamiliar, although it is possible to front an argument of the verb (if so the construction is marked by a form *l* after the fronted constituent: *Adam lə o-ikat-on muabari an*), so our Notsi speaker could have a derived SVO structure if he wanted. Whether he does or not, the verb constitutes our first real problem. For one thing, the verb stem selected decides what affixation order is possible, and what affixes are to be used (cf. Appendix III). In this case, the verb is of class I, so the speaker has to know the forms for that (again, it is unclear whether all other forms for the same pronominal category should be regarded as a source of confusion, or just those for subjects, or those for subjects as well as objects, given that both types are verbal affixes). The singular masculine subject marker further comes in two versions, future and non-future. In addition, object marking has to be considered, obligatory in a transitive construction, and a novel thing for a Notsi speaker, who would only be familiar with rather limited transitivity marking. In this case, we have a singular object, meaning that the distinction between feminine and

masculine is relevant, also an unfamiliar concern from a Notsi point of view.

In the object noun phrase there is a noun and an alienable possessive pronoun; the possessive is an unfair complication for our comparison as the Nalik phrase did not involve possession, but we may comment that the possessive pronoun contains the categories for singular masculine possessor and singular possessee (the singular masculine possessor form happens not to index the gender of the possessee), so this is quite a difficult area of Kuot grammar to master for a learner.

To sum up, the Nalik sentence requires an article (choice of four, if $\emptyset$, *na* and a specific article *ta* are included), plural (choice of 1 or none), a subject marker (choice depends on assumptions of what are confusable items), and another article for the last noun phrase.

To produce a correct Kuot sentence, the gender of the object noun has to be known and cross-referenced, the affix form and place depending on the class of the verb, which also governs the form and position of the subject cross-referencer, some of which also pay attention to tense (range of choice again depending on our assumptions). Even for producing an incorrect Kuot sentence, the differing orders of obligatory verbal affixes constitute a hurdle.

4.1 Evasion

So far so good. But we have been cheating a little by considering a ready-made sentence to match. A question that arises is how we are to think of all the optional marking for tense and aspect in the verb phrase, for instance; they add to the possible choices on a systemic level but would not always be ‘active’ or relevant in a normal speech situation. Non-obligatory items could perhaps be evaded in early stages of acquisition, and in fact I have understood that Kuot speakers conversant in Lavatbura-Lamusong (the nearest language in the Madak group which they picked up from school mates in the mixed school) get by with a reduced system of noun markers.⁵ The possibility of evasion as a mechanism in early production should therefore also be considered, but is very difficult to evaluate reliably, and attitudes among the speakers of the target language are likely to be relevant too. That is, if a learner uses no tense or aspect marking in a Nalik verb phrase even when it is expected, will it be accepted by Nalik speakers as communicatively valid? Will Notsi speakers accept learners perhaps using the plural where the trial, which is not in the other languages, had been expected? The case of Kuots reducing Lavatbura-Lamusong noun markers suggests that this is quite acceptable learner behaviour. The fact that Kuot pronominal marking on verbs, especially in verb classes II and III, is much more tightly integrated with the stem than in the Oceanic languages means

that there is no way of avoiding cross-referencing, and the fact that this is done in several different ways in turn means that you can say very little without quite a bit of knowledge of this area of grammar. It would seem that even speaking *bad* Kuot requires more extensive knowledge of the system than speaking bad, or just limited, Nalik.

4.2 Mismatches: category content and cooccurrence restrictions

Although a learner's language may have particular categories, and even code them in a similar way to the target language, there can be mismatches in application. One such item in these languages concerns the use of an inalienable possession construction. In Madak (which has inalienable forms only for singular possessors) and Notsi, it is used for some kin terms and some body parts. In Kuot, inalienable possession is used for all types of parthood, but alienable possession marking is used for kin (and in Nalik, the inalienable possessive construction has essentially fallen out of use altogether).

Another area of mismatch is the cooccurrence restrictions on markers of different grammatical categories. For example, in all the languages, the marker for future time reference has irrealis connotations, but the use of the future marker with negation operates differently, such that in Kuot there is a disassociation of future marking and negation marking, while in Nalik, the

future/irrealis marker *pe(n)* is obligatory with negation at least for younger speakers (who have reinterpreted it as a pure irrealis marker; Volker 1998: 59), while another future marker *na* can be used freely with negation. In Nalik, the durative marker may not be used with negation, while in Notsi it may, and so forth. It is not clear to me at this point how this type of mismatch would count in our difficulty scheme.

4.3 Form transparency

As mentioned above, the languages are part of a phonological alliance, sharing processes such as the lenition of voiceless stops in intervocalic position. Not all the processes and restrictions are identical in all the languages but many recur. Phonology, therefore, is not a great obstacle to interlinguistic understanding in these languages.

But Kuot and Madak have some additional processes. Madak has processes of vowel harmony that occur between stems and markers, and between concatenated markers, as well as vowel deletion in stems depending on syllabification with prefixes (e.g., *la-ban-tixin* NM-thin-woman ‘old woman’, but *la-tkin* NM-woman ‘woman’; *la-pke* ‘mud’, but *la-xam-pixe* ‘much mud’).⁶ However, these appear to be more of a hindrance to impeccable production than to comprehension, as speakers appear to be aware of the shape of the underlying form in most cases, and have no

trouble giving the unaffixed form of a noun if asked, and even Kuot speakers near the language boundary tend to know quite a bit of vocabulary from the Madak language closest to them (Lavatbura-Lamusong), and have a clear idea of the basic forms of nouns.

In Kuot, there are morpho-phonological processes that take place between stem-initial vowels and vowels of the subject prefixes in verb classes II and III. These often obscure the forms of the prefix and/or the verb stem which may make it more difficult for a learner to make sense of the system, for example /i-onəma/ is pronounced [jonəma] ‘she sits/lives’, but /u-onəma/ ‘he sits/lives’ becomes [unəma], and /pa-onəma/ ‘we (excl.) sit/live’ becomes [pon ma].

Processes such as those in Madak and Kuot conspire to obscure the form of the stem and sometimes the affixes, which is of course unhelpful to a learner.

For Kuot, the presence of morphological processes like the subtraction of many noun singular endings in plural the formation (see above), and the fact of variable affixation order for verbs is not going to help a learner: the observable representation of morphemes is not constant. From the point of

view of an Oceanic speaker, the variation associated with gender may well seem random too.

4.4 Vocabulary

Another area of potentially variable complexity is vocabulary, which should be mentioned even if not computable within the simple metric explored here.

There is far-reaching alignment in the types of meanings encoded and the scope of meaning of many words across the languages. As an example, the word normally translated as ‘kill’ is typically not result-oriented but action-oriented, implying severe beating up and such, so that an utterance like ‘he killed her so she was really upset’ is perfectly logical. This is reflected in Tok Pisin too, the structure of which is of course based on the local languages in this region. Another item similarly reflected in Tok Pisin is a word *inap* (from English ‘enough’) which means ‘enough’ as well as ‘to align’ (in the transitive form *inapim*), ‘to reach’, and ‘to be able to’, and the same polysemy is found in Kuot *puo* and *pupuo*; Notsi *pupua* means ‘enough’ and ‘be able to’; Nalik *faraxas* is ‘be same as’ and ‘be able to’; and Madak *epovo* means ‘be enough’ and ‘be able to’ (for each Oceanic language, these are the meanings I have been able to find; there may be more). There are many many similar cases, pointing to the extended contact

between speakers of the different languages over time. However, it is possible that Kuot has a richer vocabulary in the sense of having more lexical stems: speakers of Kuot and speakers of the neighbouring Madak language Lavatbura-Lamusong agree that Kuot has more words and “little meanings”. For example, while Lavatbura-Lamusong frequently derives words for fruits, tree, leaves etc. from a single lexical item, Kuot often has separate lexemes for each, and there are indications that there is more resolution also among verbs. The reason why Kuot should have more words than its neighbours is that it appears to have borrowed from several of its Oceanic neighbours and given the acquired words different functions. This information on vocabulary size is somewhat anecdotal and would need to be researched in a principled way before being admitted as support for a claim that Kuot is more difficult to acquire than its neighbours.

The influence of many languages is generally accepted as the reason for the large size of the vocabulary of English. However, in the English case there has been a levelling (simplification) of the grammar which is often attributed to the presence of second-language speakers in leading positions. It is interesting to note that such levelling appears not to have taken place in Kuot to the same extent, and perhaps the reason is that fewer outsiders, or too small a number of outsiders have learnt it.⁷

4.5 Popular judgment of difficulty

Kuot is not the only Papuan language in a setting of Austronesian languages to be judged difficult by its neighbours. At least two other cases are discussed in the literature: Anêm of New Britain and Waskia of Karkar Island, Madang Province.

Thurston (1982; 1987; 1992) writes about Anêm, a Papuan language of West New Britain which is surrounded by Oceanic languages, that “[a]ll [Austronesian]-speaking peoples in the area consider Anêm far too difficult to learn”, and he quotes Lusi speakers saying that the “Anêm language is heavy in our throats” (1982: 11). Thurston (1987), much like the above presentation of Kuot and its neighbours, gives paradigms of subject and possessive markers, and discusses stem form transparency, phonological processes, the number of grammatical distinctions made, and other factors, and convincingly shows that Anêm is more complex than its neighbours, and that this is one important reason why it has few learners. Other factors include some of the Oceanic languages being *lingue franche* with little identificational value for the original speakers, while Anêm has become a marker of its speakers’ identity and traditional culture. Yet, on the levels of syntax and semantics, all the languages of north-west New Britain have converged so that their syntactic and semantic structures are virtually identical, as a result of long association with intermarriage, shared culture,

and multilingualism as well as language shift by the original inhabitants to the languages of the Austronesian newcomers.

It may be noted that the situation of Kuot differs in that it has no emblematic function, and although semantic structures are very similar to those of neighbouring languages and most of the grammatically encoded categories are the same, it is morphologically and syntactically distinct, and it has even retained VSO constituent order in the face of very dominant SVO among its neighbours.

McSwain (1977) gives an anthropological account of the population on Karkar Island, which is divided between speakers of Papuan Waskia and Austronesian Takia. She notes that the traditional social structure of the two groups was identical, but the difference between the languages made much of. She quotes speakers talking about the difference: “‘Takia’, they said, ‘lies easy in the mouth and can be learned quickly; Waskia lies heavy and is a difficult language to master’”. These perceptions are even encoded in a myth about two brother deities from whom the island’s population descended, and from whom the Takia inherited their ‘melodious’ language and the Waskia their ‘cumbersome’ language (McSwain 1977: 4).

Interestingly, these very languages are famous for being morpho-syntactic carbon copies of one another – what Ross (e.g., 2001) has termed metatypy: the total convergence of structure but with distinct lexical and

morphological forms. In this case, the Papuan language, Waskia, was the primary inter-community language and it is Waskia's structure that has been copied by Takia. According to Ross (pers. comm.), Takia is phonologically more complex, and a case for Waskia being morphologically more complex could be made but is not very strong – yet Waskia is considered more difficult.

In the examples just shown, it seems that the perception of Anêm as difficult could be put down to complexity. But for Waskia, it is hard to argue that it is more complex than Takia, yet the perception persists. In the case of Kuot, we could well imagine that Kuot would be considered difficult on the basis of having gender alone, even if it were equivalent in complexity or even simpler than the surrounding languages in all other respects. It is clear that there is reason to be cautious regarding popular judgments of difficulty, and that these cannot always be reduced to complexity.

4.6 Extra-linguistic factors

Naturally, language structure is not all that goes into any particular learningsituation. Social and cultural factors involve items such as the status of the target language, but also attitudes towards imperfect language use by a learner, and expectations on individuals of monolingualism or multilingualism. There would also be socio-cultural differences in learning

styles between different societies. On top of these factors, individuals within the same society differ from one another. They may have different levels of motivation for learning a language, for instance depending on business interests and their personal social networks. Individuals will vary in the amount of exposure to the target language they have had, and during what period of their lives, and individuals further have varying talent for and personal interest in language learning.⁸ Each of these factors warrant research and papers and books of their own, and I just mention them here to indicate that a study of difficulty in terms of language structure is never likely to have predictive precision, although it may provide a baseline take on the learning enterprise.

Neither Kuot nor Tok Pisin enjoys any particular social status. It may be added here that the vernaculars of this area appear never to have been important for social identity, and they did not even have names until the 1950s. Imperfect language use appears largely to be tolerated, and multilingualism was common.

6. Conclusions

If we accept the premise that language complexity can be computed in an objective fashion (utopian though it may be at present), it follows that

complexity can also be compared between languages and some languages found more complex than others, in terms (minimally) of properties such as paradigm size and regularity.

Some types of complexity are likely to equal learner difficulty, perhaps especially paradigmatic irregularity, suppletion and the like, which mean that forms cannot be regularly derived but have to be learnt individually.

Other types of complexity, such as categorial distinctions, figure as difficult primarily when the learner's language does not have the same distinctions, or they are marked very differently. While the different verb affixation orders of Kuot would presumably always be a challenge to a learner, the fact of object cross-referencing is likely to be less of a problem for someone whose own language has object cross-referencing; it would be a matter of remapping a familiar function rather than acquiring an entirely new kind of grammatical behaviour.

As we have seen, e.g., the syntax of Kuot verb phrases is not necessarily more complex than in the neighbouring languages, but there is still reason to assume that a speaker of a neighbouring language would find even this area a lot more difficult to learn than in any of the other neighbouring languages, due to the lack of overlap of marking structures.

Even if we accept that the structures of two languages can be compared in such a way as to reveal wherein the difficulties for a learner from a particular linguistic background will lie, we will still have only a rather crude tool for gauging the difficulty in any actual learning situation, as in any concrete situation a variety of non-negligible extralinguistic factors come into play, including social, cultural and individual variables. With the societies considered being socio-culturally so homogenous, it would be hard to explain the more rapid decline of Kuot compared with its neighbours without reference to the language itself. I hope to have shown that Kuot is more complex in several areas, but, above all, difficult for speakers of neighbouring Oceanic languages.

It is possible that the Kuots have been just that little bit more used to switching into other languages to accommodate non-Kuots, and that this habit is enough to tip the balance such that the language is about to be abandoned altogether. The Oceanic languages surrounding Kuot are by no means safe from the threat of extinction, but so far children in most places appear to learn them, unlike Kuot children.

Acknowledgements

I wish to thank one anonymous referee, the editors and Wouter Kusters for feedback which has much improved this paper. I also want to thank Lee

Erickson and Bob Lee for letting me use their unpublished data on Notsi and Madak respectively; the latter also for discussion on the language situation. Gunnar Eriksson has read draft versions of this paper and provided useful discussion. Initial parts of this work, as part of the European Science Foundation EUROCORES Programme OMLL, was supported by funds from Vetenskapsrådet and the EC Sixth Framework Programme under Contract no. ERAS-CT-2003-980409.

Glosses used in examples

APOSS	alienable possessive
ART	article
CN	continuous aspect
DEM	demonstrative
F	feminine
INDEF	indefinite
M	masculine
NFUT	non-future
NM	noun marker
NSG	non-singular
OBJ	object
PL	plural
SG	singular
SUBJ	subject

References

- Dahl, Ö. 2004. *The growth and maintenance of linguistic complexity*.
Amsterdam/Philadelphia: John Benjamins.
- Erickson, L. and Erickson, L. 1991. *Grammar Essentials of the Nochi
Language of New Ireland Province*. ms.
- Gordon, R.G. Jr. (ed.). 2005. *Ethnologue: Languages of the World*,
Fifteenth edition. Dallas, TX: SIL International. Online
version: <http://www.ethnologue.com/>.
- Lee, R. 1978. *Madak Grammar Essentials*. ms.
- Lee, R. 1989. The Madak verb phrase. *Language and Linguistics in
Melanesia* 20(1–2): 65–114.
- Lee, R. n.d. *The Madak Noun Phrase*. ms.
- Lindström, E. 2002. Topics in the grammar of Kuot – a non-Austronesian
language of New Ireland, Papua New Guinea. PhD Dissertation,
Stockholm University.
- Lindström, E. in prep. Phonological convergence in new ireland: Kuot and
its austronesian neighbours.
- McSwain, R. 1977. *The past and future people – Tradition and change on a
New Guinea island*. Melbourne: Oxford University Press.
- Miestamo, M. 2005 [/in press]. On the Feasibility of Complexity Metrics.
In: *Finest Linguistics. Proceedings of the Annual Finnish and Estonian*

Conference of Linguistics, K. Kerge and M-M. Sepper (eds). Tallinn, May 6-7, 2004. Tallinn: TLÜ.

Ross, M. 1994. Areal phonological features in north central New Ireland. In *Language Contact and Change in the Austronesian World* [Trends in Linguistics, Studies and Monographs 77], T. Dutton and D.T. Tryon (eds.), 551–572. Berlin: Mouton de Gruyter.

Ross, M. 2001. Contact-Induced Change in Oceanic Languages in North-West Melanesia. In: *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics*, A.Y. Aikhenvald and R.M.W. Dixon (eds). Oxford/NY: Oxford University Press.

Spriggs, M. 1997. *The Island Melanesians*. Oxford: Blackwell.

Thurston, W.R. 1982. *A comparative study in Anêm and Lusi*. Canberra: Pacific Linguistics.

Thurston, W.R. 1987. *Processes of change in the languages of north-western New Britain*. Canberra: Pacific Linguistics.

Thurston, W.R. 1992. Sociolinguistic typology and other factors effecting change in north-western New Britain, Papua New Guinea. In *Culture change, language change – case studies from Melanesia*, T. Dutton (ed), 123–139. Canberra: Pacific Linguistics.

Volker, C.A. 1998. *The Nalik Language of New Ireland, Papua New Guinea*. New York: Peter Lang.

¹ The reference of the term Madak can be confusing, as it is the name of a small language family ‘the Madak languages’, as well as one of its members, the languages being Lavatbura-Lamusong, and Madak. Even within each of these, there is much dialectal variation, and Lee (1978, 1989) cautiously restricts the scope of his observations to the Malom variety, which, unfortunately for our purposes, is non-adjacent to Kuot.

² The claim in the Ethnologue that Kuot is “vigorous on the west coast” is not supported by my own observations. In general, it would have been desirable to have more detailed sociolinguistic information on the Oceanic languages discussed here, but unfortunately such information is not available to me at the time of writing.

³ Cf. Dahl’s concept of choice structure (2004: 46–50).

⁴ For ‘he’ in this scenario, read ‘he or she’.

⁵ I do not speak or understand Lavatbura-Lamusong but have heard particular speakers of Kuot speaking it seemingly competently. However, discussing it later, they were unaware of meaning differences beyond grammatical number between nouns prefixed in different ways (I knew of these from interviewing Lavatbura-Lamusong speakers).

⁶ In Madak orthography, <x> stands for the voiced velar fricative [ɣ], which is an intervocalic allophone of /k/; as mentioned above this is a process common to several of these languages (cf. Ross 1994; Lindström in prep.)

⁷ An area of grammar that may be seen as a possible exception is the verb classes and adjectives (cf. Appendix III): the only open verb class is that of the least morphological complication, likely derived from a noun+light verb construction. However, there are several hundred stems still in use of classes II, III and adjectives, including many basic verbs, auxiliaries and serializing verbs. The loss of productivity of these classes could tentatively be accounted for by the presence of non-native speakers, and it may be noted that today’s semi-speaker young Kuots typically have trouble with less common verbs of those classes.

⁸ I fear that mentioning individual differences in talent for language acquisition is as politically incorrect as it was until recently to imply that languages may differ in complexity. However, I consider it an empirical fact that people differ in this respect, both in Western society and in the Papua New Guinea society where I have worked.

Appendix I – Verb phrase

The Nalik verb phrase

S	COUNTERFACT	TNS/ASP	REC	DUR/NEG/ LOC	CAUS/ REDUP	V	INC/SER	DL/PC	TRANS	FOC	COMPL
	<i>lek</i>	ANT <i>tabung</i> FUT <i>na</i> INCEPT <i>vala</i> HAB <i>runa</i>	<i>vara</i> = <i>vavur</i>	DUR <i>i, -t</i> NEG <i>pe(n)</i> LOC <i>su, o</i>	CAUS <i>fa</i>		[inc. O; DO; ser V]	DL: <i>-(y)a</i> PC: <i>-(t)al</i>	<i>-ing</i>	<i>ang</i>	<i>faanong</i>

The Notsi verb phrase

S	TNS+	PERF DUR 'CONT' lo SEQ +	ASP+	NEG	MOOD+	REC	V	TRANS	ADV	MOTION	TRANS	LIM	AUG
	FUT <i>ba</i> PST <i>ti-</i> CERT <i>ta</i>		<i>se</i> <i>mager</i> <i>lo</i> <i>ga</i> +	<i>kap</i>	IMPER <i>na</i> APPR <i>soro</i> PERM <i>le</i> INDEF <i>ka, mara</i> HAB <i>tigiri</i> ABIL <i>pupua</i> INTS <i>suuk</i> INTS <i>mun</i>	<i>e-</i>		<i>-i</i>		<i>la</i> 'go'	<i>ngin</i>	<i>mu</i>	<i>buka</i>

(Causative *a-* occurs after mood but it is not clear in what relation to reciprocal).

The Madak verb phrase

S	IRR	TNS	NEG	ASP	X	CONT	V	MODIF
	<i>gi-</i>	YPST <i>ta</i> RPST <i>ga</i> IFUT <i>na</i> RFUT <i>ba</i> CFUT <i>naba</i> PRES/HAB Ø	-xo- (<i>ta</i>) *	PRG- <i>a-</i> , - <i>ra-</i> **	CAUS - <i>vaxa-</i> REC - <i>e-</i> TOT - <i>va-</i>	[redup]		[verb semi-V noun adverbs]

* *ta* after *gi-*. Phonological interaction with tense *ta*, *ga*, *ba* → *toxoxo* etc.

** *a* is present, *ra* vague past

Madak allows for only three markers to be present at any one time.

The Kuot verb phrase

FUT/NEG	•	AUX	SER	•	V	•
FUT <i>e(ba)</i>		HAB - <i>me</i>	‘go’ - <i>la</i>		*	
NEG <i>təle</i>		HAB <i>buat</i>	‘come’ <i>mu-o</i>			
FNEG <i>təla</i>		DESID - <i>ga</i>	‘come’ - <i>op</i>			
		DESID <i>nəmo</i> (?) - <i>ma</i>				

Dots (•) mark common positions for adverbs

* Verb affix order depends on verb class and grammatical person

2nd pos	free pos
ASP/ADV PTCL	CONT
ASP - <i>arə</i> ‘just’ - <i>it</i> ‘little’ - <i>arom</i> + ‘yet’ <i>ka</i> EMPH <i>kan</i> ‘too’ <i>gət</i> ‘now’ <i>bət</i>	<i>mən</i>

Abbreviations used in Appendix I: + ‘and more’; ABIL ability; ADV adverbial; ANT anterior; APPR apprehensive; ASP aspect; AUG augmentative; AUX auxiliary; CAUS causative; CERT certainty; CFUT certain future; COMPL complete; CONT continuous; COUNTERFACT counterfactual; DESID desiderative; DL dual number; DUR durative; EMPH emphatic; FNEG future negation; FOC focus; FUT future; HAB habitual; IFUT immediate future; IMPER imperative; INC/SER incorporated DO / serial verb; INCEPT inceptive; INDEF indefinite; INTS intensive; IRR irrealis; LIM limiter; LOC locative; MODIF modifier; NEG negation; PC paucal; PERF perfective ; PERM permissive; PRES present; PST past; REC reciprocal; REDUP reduplication; RFUT remote future; RPPST remote past; SEQ sequence; SER serial (verb); TNS tense; TOT totality; TRANS transitive; X various; YPST yesterday past

Appendix II – Pronominal forms

Forms which have already occurred in the same person-number function have been grayed.

Nalik pronominal forms

		DL: -a			
	indep	S mkr	PC: -(mt)al	alien poss	inal poss
SG	1 <i>ni</i>	<i>ga</i>		<i>surago</i>	<i>-nagu, -nugu</i>
	2 <i>nu</i>	<i>gu</i>	+	<i>sunum</i>	<i>-num</i>
	3 <i>naan</i>	<i>ka, a, na</i>		<i>si-</i>	<i>-na</i>
NON	1n <i>di</i>	<i>di(a)</i>	+	<i>si-</i>	<i>-di</i>
-SG	1x <i>maam</i>	<i>madi</i>	+	<i>si-</i>	<i>-maam</i>
	2 <i>nim</i>	DL/PC**: <i>gu</i>	+	***	DL <i>-numa</i>
		PL: <i>nagu</i>			PC <i>-numtal</i>
3	<i>na(a)nde,</i>	<i>di(a),</i>	+	<i>si-</i>	PL <i>-nim</i>
	<i>na(a)ndi,</i>	<i>(ka)</i>			<i>-naande,</i>
	<i>na(a)nda*</i>				<i>-naandi,</i>
					<i>-naanda</i>

* dialectal and generational variation

** DL and PC are unusual; if used they have to go both on pronoun and V; 2nd person can have sg or nsg as base

*** DL in all persons by adding *si-* and *-(y)a* to PL inal; PC in all persons by adding *si-* and *-yal* or *-tal*

Notsi pronominal forms

	S/indep	O/prep	alien poss	inal poss	ben-n-food	ben-food
SG	1 <i>ya</i>	<i>ya</i>	<i>mugu</i>	N-gV	<i>na-go</i>	<i>ma-ga-ka</i>

	2	<i>u</i>	<i>yu</i>	<i>nuum</i>	N-V _m	<i>nuum-ka</i>	<i>maam-ka</i>
	3	<i>a-/Ø</i>	Ø	<i>kan</i>	N-n	<i>ka-ka</i>	<i>ma-ka</i>
DL	1n	<i>gita</i>	<i>gita</i>	<i>kida</i>	+ N-n	<i>mida</i>	–
	1x	<i>gelu</i>	<i>gelu-lu</i>	<i>na-gelu</i>	+ N-n	<i>na-gelu-ka</i>	<i>ma-gelu-ka</i>
	2	<i>gulu</i>	<i>gulu-lu</i>	<i>na-gulu</i>	+ N-n	<i>na-gulu-ka</i>	<i>ma-gulu-ka</i>
	3	<i>delu</i>	<i>delu-lu</i>	<i>ka-delu</i>	+ N-n	<i>ka-delu-ka</i>	<i>ma-delu-ka</i>
TL	1n	<i>gita-tuul</i>	<i>gita-tuul</i>	<i>kida-tuul</i>	+ N-n	<i>mida-tuul</i>	–
	1x	<i>getu</i>	<i>getu-tuul</i>	<i>na-getu</i>	+ N-n	<i>na-getu-ka</i>	<i>ma-getu-ka</i>
	2	<i>gutu</i>	<i>gutu-tuul</i>	<i>na-gutu</i>	+ N-n	<i>na-gutu-ka</i>	<i>ma-gutu-ka</i>
	3	<i>detu</i>	<i>detu-tuul</i>	<i>ka-detu</i>	+ N-n	<i>ka-detu-ka</i>	<i>ma-detu-ka</i>
PL	1n	<i>gita'</i>	<i>gita'</i>	<i>kida'</i>	+ N-n	<i>mida'</i>	–
	1x	<i>geem</i>	<i>geem</i>	<i>na-geem</i>	+ N-n	<i>na-geem-ka</i>	<i>ma-geem-ka</i>
	2	<i>gim</i>	<i>gim</i>	<i>na-gim</i>	+ N-n	<i>na-gim-ka</i>	<i>ma-gim-ka</i>
	3	<i>di</i>	<i>di</i>	<i>ka-di</i>	+ N-n	<i>ka-di-ka</i>	<i>ma-di-ka</i>

Inalienable possessives in non-singular numbers are formed in two ways:

1. alienable + possessed N + -n, or 2. possessed N + -n + object pronoun.

Madak pronominal forms

	Indep S (<i>ne-</i>)	S mkr	Indep O	at- poss, prep:loc,tim	mi- comit, instr	ti- benef, goal etc	inal poss
SG	1	<i>nia</i>	<i>ia</i>	<i>-arak</i>	<i>-nia</i>	<i>-a</i>	<i>-k</i>
	2	<i>nu</i>	<i>u</i>	<i>-aram</i>	<i>-nu</i>	<i>-u</i>	<i>-m</i>
	3	<i>ni</i>	<i>i</i>	<i>-eren</i>	<i>-n</i>	<i>-n</i>	<i>-n (-no/-na/-on/)</i>
DL	1n	<i>da</i>	<i>da</i>	<i>-da</i>	<i>-da</i>	<i>-da</i>	<i>-da</i>
	1x	<i>ma'</i>	<i>nama</i>	<i>-nama</i>	<i>-nama</i>	<i>-nama</i>	<i>-nama</i>

2	<i>mu</i>	<i>mu</i>	<i>numu</i>	<i>-numu</i>	<i>-numu</i>	<i>-numu</i>
3	<i>du</i>	<i>du</i>	<i>du</i>	<i>-du</i>	<i>-du</i>	<i>-du</i>
PL 1n	<i>dik</i>	<i>ta</i>	<i>dik</i>	<i>-dik</i>	<i>-dik</i>	<i>-dik</i>
1x	<i>ma</i>	<i>ma</i>	<i>nama</i>	<i>-nama</i>	<i>-nama</i>	<i>-nama</i>
2	<i>mi</i>	<i>mi</i>	<i>nimi</i>	<i>-nimi</i>	<i>-nimi</i>	<i>-nimi</i>
3	<i>di/ti</i>	<i>di**</i>	<i>di</i>	<i>-di</i>	<i>-di</i>	<i>-di</i>

* omitted before tense marker

** omitted after *nedi*

Kuot pronominal forms

	Pers		Dem; attr.	Subj aff Vcl.II,III, n-fut/fut	Inal poss, prep.	Obj prep.	Adj.	Obj suff. Vcl.IIa 3rd, n-fut/fut	Subj encl., Vcl.I, n-fut/fut
	num	pers	relator						
SG 1	<i>turuo</i>	–	–	<i>tu/ta</i>	<i>tuo</i>	<i>to</i>	<i>to-</i>	–	<i>tuy/taŋ</i>
2	<i>nunuo</i>	–	–	<i>nu/na</i>	<i>nuo</i>	<i>no</i>	<i>no-</i>	–	<i>nuŋ/naŋ</i>
3m	–	<i>i(-)</i>	–	<i>u/a</i>	<i>a</i>	<i>a</i>	<i>-i</i>	<i>-a/-ŋ</i>	<i>oŋ/aŋ</i>
3f	–	<i>u(-)</i>	–	<i>i</i>	<i>o</i>	<i>o (u)</i>	<i>-u</i>	<i>-o/-ŋ</i>	<i>ieŋ</i>
DL 1n	<i>bibi</i>	–	–	<i>bi</i>	<i>bi</i>	<i>bi</i>	<i>bi-</i>	–	<i>biŋ</i>
1x	<i>i</i>	–	–	<i>i</i>	<i>i</i>	<i>i</i>	<i>i-</i>	–	<i>iŋ</i>
2	<i>mame</i>	–	–	<i>ma</i>	<i>me</i>	<i>me</i>	<i>ma--n</i>	–	<i>maŋ</i>
3	–	<i>li-</i>	–	<i>li</i>	<i>li(e)</i>	<i>le</i>	<i>-n</i>	<i>-an/-ŋan</i>	<i>lioŋ</i>
PL 1n	<i>bubuo</i>	–	–	<i>bu</i>	<i>buo</i>	<i>bu</i>	<i>bu--m</i>	–	<i>buoŋ</i>
1x	<i>papa</i>	–	–	<i>pa</i>	<i>pa</i>	<i>pa</i>	<i>pa--m</i>	–	<i>paŋ</i>
2	<i>mimi</i>	–	–	<i>mi</i>	<i>mi</i>	<i>mi</i>	<i>mi--m</i>	–	<i>miŋ</i>
3	–	<i>mi(-)</i>	–	<i>me</i>	<i>ma</i>	<i>ma</i>	<i>-m</i>	<i>-am/-m</i>	<i>meŋ</i>

num	pers	Alienable possessives			
		masc	fem	dual	plural
SG	1	<i>tuay</i>	<i>tuy</i>	<i>tuayan</i>	<i>tuam</i>
	2	<i>nuay</i>	<i>nuŋ</i>	<i>nuayan</i>	<i>nuam</i>
	3m		<i>ay</i>	<i>ayan</i>	<i>am</i>
	3f	<i>iaŋ</i>	<i>ieŋ</i>	<i>iaŋan</i>	<i>iam</i>
DL	1n		<i>biŋ</i>	<i>biŋan</i>	<i>bim</i>
	1x		<i>iŋ</i>	<i>iŋan</i>	<i>im</i>
	2		<i>meŋ</i>	<i>meŋan</i>	<i>mem</i>
	3	<i>liaŋ</i>	<i>lioŋ</i>	<i>liaŋan</i>	<i>liam</i>
PL	1n	<i>buay</i>	<i>buoŋ</i>	<i>buayan</i>	<i>buam</i>
	1x		<i>pay</i>	<i>paŋan</i>	<i>pam</i>
	2		<i>miŋ</i>	<i>miŋan</i>	<i>mim</i>
	3	<i>meiaŋ</i>	<i>meioŋ</i>	<i>meiaŋan</i>	<i>meiam</i>

Appendix III – Kuot verbs and adjectives, and their affixation orders

Verb class	Intransitive	Transitive
I. subj.enclitic obj.prefix him'	<i>pasei-oŋ</i> talk-3mSUBJ 'he talks'	<i>a-pasei-oŋ</i> 3mOBJ-talk-3mSUBJ 'he talks of him/it(m)' / 'he tells him'
IIa. subj.prefix obj.suffix in 3 obj.prefix in 1/2	<i>u-libə</i> 3mSUBJ-cry 'he cries'	<i>u-alibə-o</i> 3mSUBJ-cry.over-3fOBJ 'he cries over it(f)/her [mourns]' <i>to-u-alibə</i> 1sOBJ-3mSUBJ-cry.over 'he cries over me'
IIb. subj.prefix obj.prefix	<i>u-lo</i> 3mSUBJ-talk 'he talks'	<i>a-u-lo</i> 3mOBJ-3mSUBJ-talk 'he tells him'
III. subj."infix" obj.prefix	<i>uan-u-lə</i> wait-3mS-stm2 'han väntar'	<i>a-uan-u-lə</i> 3mOBJ-wait.for-3mSUBJ-stm2 'he waits for him(/it.m)'
adj. subj.suffix i 3 subj.prefix i 1/2	<i>mukə-u</i> pregnant-3f 'she is pregnant' <i>to-mukə-i</i> 1SG-pregnant-SG 'I am pregnant'	— —