Discontinuous noun phrases in Ngkolmpu

Matthew J Carroll

Abstract It has been claimed that Ngkolmpu (Yam, Papuan) displays discontinuous noun phrases (Donohue 2011). However, careful textual analysis of a corpus of naturalistic language reveals that, in practice, this is highly restricted. The data shows two relatively rare constructions which give rise to limited discontinuous structures. The first is an afterthought construction involving a full co-referential nominal constituent adjacent to the clause. This co-referential constituent is both syntactically and phonetically distinct from the main utterance. The other involves a topic marking demonstrative encliticised to verb at the right edge of the clause interacting with general information-structural conditions on word order. This is the only true discontinuity found in the corpus and is restricted to demonstratives only. This paper clarifies a claim in the literature about the empirical facts of a specific language, Ngkolmpu, and adds a nuanced discussion of nominal discontinuity in a language of New Guinea.

1   Introduction

Ngkolmpu, a Yam language spoken in the extreme south-east of the Indonesian province of Papua, has been claimed to display discontinuous noun phrases (Donohue 2011).[1] Crucial to Donohue's claims are examples such as in (1) which he argues involve a single nominal constituent, highlighted here in bold, which is scrambled throughout the clause.

---

[1] Donohue (2011) labels this language Kanum following a group level ethnonym although clarifies that he is referring to Ngkontar Ngkolmpu. As there are a number of languages that could equally be labelled Kanum, I use the more specific name of Ngkolmpu in this paper. A discussion of the various Kanum languages is presented in Section 1.1.

(1) **ntop-w**   mo   **piengku**   sr\eyerk/nt   **irye-w**
     big-SG.ERG   wallaby   **DIST.ERG** SG>3.FUT.DUR\stalk   **man-SG.ERG**

'The big man will stalk a wallaby[2].'   (Donohue 2011: 505)

However, careful study of textual data shows limited evidence for discontinuous noun phrases. Data from a sample of 1500 utterances drawn from transcribed speech collected during field work from 2012 to 2017 contains just 47 examples which show potential discontinuity. I argue that these examples are the result of up to two distinct constructions only one of which I argue is a case of true discontinuity in the sense of (Schultze-Berndt & Simard 2012). The first of these constructions is an afterthought construction (Averintseva-Klisch 2008); this involves a second co-referential nominal constituent following the main utterance. This gives rise to some structures which appear discontinuous but syntactic and phonetic evidence shows these are distinct constituents rather than a single discontinuous unit. The second involves the encliticisation of the demonstrative following the verb rather than occurring with the noun phrase. Thus we have some restricted discontinuity between noun phrases and demonstratives only. The interaction of these constructions with word order variation according to information structure further complicates the picture.

Beyond Donohue (2011), this paper draws on other studies on discontinuous noun phrases in Australian languages which examine textual data with insights from both information-structural and intonational analyses (Schultze-Berndt & Simard 2012). Additionally, it adds to the general typology of configurationality in Papuan languages (Donohue 2005).

This paper is structured as follows. In the remainder of the introduction (§1.1), I give an overview of the Ngkolmpu language. Section 2 is an overview of nominal syntax in Ngkolmpu which provides the background for the corpus study. With

---

[2] I apply a slightly different orthography to that of Donohue in this paper and follow the practice in Carroll (2016). The examples from Donohue (2011) have been transliterated into the orthography used here.

the basics of nominal syntax established, I summarise the relevant claims made in Donohue (2011) (§2.3). Section 3 is a corpus study of discontinuous structures in Ngkolmpu; it starts with a discussion of the methods and presents the two relevant constructions, describes their structure, properties and function as determined from the corpus study. Section 4 presents some reanalysis of Donohue's data in light of the study with some additional elicited examples. The paper concludes in Section 5.

1.1   The Ngkolmpu Language

Ngkolmpu is the westernmost language of the Yam family, one of the primary language families of New Guinea (Evans et al. 2017). Figure 1 is a map of the main locations where Yam languages are spoken. As is shown in the map, Ngkolmpu is a member of the Tonda-Kanum subgroup. The primary reference for the language is Carroll (2016) and a very brief discussion of the issues raised in this paper are discussed in that reference. Previous studies have also called this language Kanum (see footnote 1) however this term can refer to any of the Tonda-Kanum varieties spoken in Indonesia (Carroll 2016) and so the more specific name, *Ngkolmpu*, is preferred here.

There are two varieties of Ngkolmpu, Ngkontar Ngkolmpu and Bedi Ngkolmpu. Ngkontar Ngkolmpu is spoken by around 150 speakers in the village of Yanggandur; there is also a number of speakers living in the near by village of Rawa Biru, primarily patrilocal-residing women from Yanggandur and their families. I estimate there are at least an additional 50 speakers living in Rawa Biru based on observations made in January 2017. In Yanggandur, Ngkolmpu is used as a day-to-day language of communication along with Taemer (another Yam language) and the more dominant local varieties of Indonesian. Children are no longer learning the language and the youngest speakers are in their thirties. The other variety known as Bedi (also Baedi) Ngkolmpu was traditionally spoken in the coastal area around

Figure 1: Locations of the Yam languages.

Onggaya; it is no longer a living language and a visit in 2017 confirmed six people who remembered the language; the youngest of which is in their sixties, none of whom used the language on a regular basis.

All the data in this paper is of Ngkontar Ngkolmpu with speakers from Yanggandur. This is the same location as where Donohue (2011) collected his data. People in Yanggandur report that Donohue worked with speakers one generation older than I worked with and who are now mostly deceased. Differences in his data to my own may represent be inter-speaker or inter-generational differences, however such explanations are purely speculative. The recordings that make up my corpus are available on the PARADISEC archive (Carroll 2012). Examples indicate the file name of the recording and segment number from the text indicated by the square brackets. If no reference is given, the example was constructed for illustrative purposes and confirmed with at least two native speakers.

The language marks grammatical roles through both flagging (case marking) and indexing on the verb. Case is marked on an ergative/absolutive basis (2). Verbal indexing is more complex marking up to two arguments labelled as actor and undergoer. The details of the alignment system of agreement is beyond the scope of this paper and these can be thought of as subject and object. Verbal morphology is complex with a strong tendency to distribute exponence across both prefixes and suffixes (Carroll 2016). As such, inflectional material is separated from lexical material with slashes '\' and '/' and a unified gloss is provided for all the inflection material together. The symbol '>' is used to indicate semantic roles of arguments; so that the verb gloss in (2a), SG>3.HOD.PFV, would be read as 'singular actor *acting on* third person undergoer in the hodiernal past tense and perfective aspect.'

(2)   a.   piengku   pi   s\rampu/i
             3SG.ERG   3.ABS   SG>3.HOD.PFV\bite

             'S/he bit him/them.'

        b.   pi   t\awance/i
             3.ABS   SG.HOD.PFV\fell

             'She/he fell.'

The default word order for transitive clauses is AOV as exemplified in (2a), however all word orders are possible given the correct pragmatic or information-structure contexts. Topical or given arguments are typically omitted in discourse and the most common orders are V, OV or AV clauses (3).

(3)   a.   moro sw\mark/   pnm   mens-nm
             HOW   SG>3.RCT\burn/   DIST.INS   fire-INS

             '(I) cooked (the fish) with the fire.'   [001-01 014][3]

---

[3] These numbers are a reference to the corpus. For more details see Section 3. Any and all examples which lack a reference number or citation come from the author's fieldnotes

Word Order	Count
OV	38
V	34
AV	9
AOV	5
VA	5
VAO	2
OVA	2
VO	2
AVO	1
VOA	1
OAV	1
Total	**100**

Table 1: Count of word orders from a sample of 100 transitive clauses.

    b.   mens su\loka/i
           fire   NSG>3.RCT\make.fire/

       '(We) built a fire.'   [001-01 005]

    c.   pinta   n\saemingke/e
           3NSG.ERG 3NSG>1NSG.PRS\teach/

       'They are teaching (us).'   [003-01 007]

The word order of a randomly selected sample of 100 transitive clauses is presented in Table 1. The most common orders are OV and V. The four most common orders correspond to a general principle of AOV word order, i.e. arguments occur before the verb and A occurs before O when present. The results also confirm that all 11 possible word orders are attested; this point is important for the discussion of discontinuity later.

2 Basic nominal syntax

Noun phrase syntax is briefly covered in Donohue (2011). This is discussed in more detail in Carroll (2016). The analysis in both sources is largely the same. The relevant facts for this study are reproduced here.

Nominal constituents are either a DP or NP. The DP consists of a case marked demonstrative and a case marked NP. Typically the demonstrative follows the NP although may precede the NP when focused. The NP is strictly right headed with modifiers preceding the head noun. Bare NPs, i.e. NPs without a demonstrative, are common. NPs with an elided head are also possible; in these cases the right-most element of the NP becomes the host for the case marker.

Ngkolmpu has 11 distinct case forms in the pronouns and demonstratives and 10 for standard nouns. Nominal inflection is highly regular; these is no allomorphy beyond general phonological processes and there are no inflectional classes or nominal gender beyond natural sex. As discussed in Section 1.1, core cases are marked on an ergative alignment. Ergative is the only nominal case which marks number and only for animates. Nominal case marking makes no formal distinction between dative and possessive, but these are distinguished in the pronouns and as such are treated separately here. Additionally, many non-core cases mark different roles if they are used in adnominal, relational, or subordinating functions;[4] a discussion of this is beyond the scope of this paper. The case forms and their labels are presented in Table 2.

2.1 Demonstrative/Determiner phrases

A demonstrative phrase, by default, consists of a case marked NP followed by a case marked demonstrative (4).

---

[4] This is very similar to what has been described for Australian languages by Dench & Evans (1988).

Case	Ending	Example
Absolutive	Ø	*krar*
Ergative SG	-w	*krarw*
Ergative NSG	-ya	*krarya*
Dative	-en	*kraren*
Possessive	-en	*kraren*
Instrumental	-nm	*krarnm*
Causal	-wa	*krarwa*
Purposive	-t	*krart*
Associative	-wi	*krarwi*
Locative	-ni	*krarni*
Ablative	-mpa	*krarmpa*
Allative	-ngke	*krarngke*

Table 2: Cases exemplified for *krar* 'dog'.

(4) [[yuow ntop krar-ya]ₙₚ   piengku]ᴅₚ ngko
      three big   dog.NSG.ERG DIST.ERG   1SG.ABS
     b\rampu/i
     NSG>1SG.HOD.PFV\bit/

     'Those big three dogs bit me.'

Occasionally the demonstrative precedes the NP (5). This is restricted to new information.

(5) atka   moi   owne   [ᴅₚ piengku [ₙₚ pipa-w]]
     water where SG.PRS.DUR\drink   DIST.ERG   pipe-SG.ERG
     Merauke
     Merauke.'

     'Where the pipe sucks water for Merauke.'   [010-001 027]

Demonstratives may function pronominally, in which case they occur without an NP. Example (6a) shows the demonstrative alone. The demonstrative may also occur with a full NP as shown in (6b).

(6)   a.   [ngkiengku]_{DP} bori s\en/ngk   pien
                  PROX.ERG   BG   SG>3.PST.PFV\give/ 3.DAT

        'This (one) gave (it) to him'   [016-02 453-7]

        b.   [[irepe-u]_{NP}   ngkiengku]_{DP} bori s\en/ngk
                  person-SG.ERG PROX.ERG   BG   SG>3.PST.PFV\give/
               pien
               3.DAT

        'This person gave (it) to him'

2.2   Noun phrases

Noun phrases are strictly right headed. A nominal head can be preceded by numeral and modifier in that order. The head is marked for case if non-absolutive. The basic noun phrase template is presented in (7).

(7)   [(NUM) (MOD) N]

The order within a NP is strict with no alternate orderings attested. The modifier slot may be filled with an adjective (8), derived adjective (9), noun (10) or infinitive (11). The demonstrative may never occur within the NP.

(8)   yuow ntop krar   (10)   yuow mwa   krar
         three big   dogs   three house dogs
         'Three big dogs'   'Three house dogs'

(9)   yuow kence-wa   krar   (11)   yuow para   krar
         three heaviness-ADJ dogs   three hunt.INF dogs
         'Three heavy dogs'   'Three hunting dogs'

It is common for bare NPs to occur without a demonstrative (12).

(12)   krar-ya   mo   poi   su\ruo/rntne
         dog-PL.ERG wallaby DIST.LOC NSG>3.RMT.IPFV\chase/

         'The dogs chased a wallaby there.'   [009-01 006]

It is possible to elide the head noun. In these cases the rightmost element is marked for case. Examples (13a) show a derived adjective as the only nominal in the phrase and (13b) shows similarly shows an underived adjective in this position.

(13) a. pla-wa-w   n\mk/w   ngkei
          white-ADJ-SG.ERG   SG>1NSG.RMT.DUR\gather.EX   here

          'The whites gathered (us) here.'

      b. ntop-w   mo   pi   y\merk/i
          big-SG.ERG   wallaby   DIST.ABS   SG>3.HOD.DUR\follow

          'The big (one) followed that wallaby.'

2.3   Discontinuity in Donohue (2011)

Based on a suite of elicited examples, Donohue (2011) argues that Ngkolmpu may display discontinuous noun phrases for noun phrases marked with overt structural case, i.e. a case marking a core argument.

The basic claim comes from a series of examples which display apparently discontinuous nominal constituents reproduced in (14) (Donohue 2011: 505). Note that the bracketing in these sentences corresponds to the analysis in Donohue (2011) rather than the analysis presented in Section 4.

(14) a. [[ntop yrye-w]$_{NP}$ piengku]$_{DP}$ sr\eyerk/nt
             big   man-SG.ERG DIST.ERG   SG>3.FUT.DUR\stalk/

      b. [ntop yrye-w]$_{NP}$ sr\eyerk/nt   [piengku]$_{DP}$
             big   man-SG.ERG SG>3.FUT.DUR\stalk/   DIST.ERG

      c. [[yrye-w]$_{NP_i}$ piengku]$_{DP}$ sr\eyerk/nt
             man-SG.ERG DIST.ERG   SG>3.FUT.DUR\stalk/
          [ntop-w]$_{NP_i}$
             big-SG.ERG

      d. [[ntop-w]$_{NP_i}$ piengku]$_{DP}$ sr\eyerk/nt
             big-SG.ERG DIST.ERG   SG>3.FUT.DUR\stalk/
          [yrye-w]$_{NP_i}$
             man-SG.ERG

e. [ntop-w]$_{\text{NP}_i}$ mo   [piengku]$_{\text{DP}}$ sr\eyerk/nt
       big-SG.ERG wallaby   DIST.ERG   SG>3.FUT.DUR\stalk/
    [yrye-w]$_{\text{NP}_i}$
       man-SG.ERG

'That big man will stalk it/a wallaby.'

Example (14a) represents the default nominal word order: a DP constituent consisting of an NP followed by a demonstrative; both the right-most element of the NP and the demonstrative are marked for case. The remainder of the examples involve some element or another separated from the DP in clause final position. Example (14e) shows all three elements of the DP occurring non-adjacently. Donohue (2011) argues that these are the result of a process of scrambling. We shall return to this claim and set of examples in Section 4. In that section, I propose that examples (14b)-(14d) are able to accounted for as afterthought constructions although no prosodic data is available to confirm this. Example (14e) can potentially be accounted for as the interaction between topic marking and an afterthought construction.

Donohue (2011) goes on to argue that this construction is only possible with nominals marked with an overt structural case. We shall see in Section 4 that this is more complex; with examples occurring in both the absolutive, which is not-overt, and instrumental, which is not a structural case.

3   Discontinuity in the corpus

This study is based on a sample of 1500 utterances of semi-naturalistic data. This sample is drawn from a larger corpus collected over 10 months of fieldwork in Yanggandur between 2012 and 2016. The sample consists of a variety of text types, including narratives, recounts, procedural texts, picture tasks and conversations. Texts were collected from 4 speakers, 2 male and 2 female aged from their mid-40s

to mid-50s.[5] The data is time-aligned to audio recordings and glossed to the word level.

From this sample, I extracted all utterances which contained two non-contiguous nominal elements of the same case including absolutive. Examples which involved two distinct referents were excluded. This left a total 46 utterances which showed potentially discontinuous noun phrases. From these 46, 12 were what I label an afterthought construction and 34 involved the topic clitic demonstrative potentially interacting with general processes affecting word order.

3.1 Afterthought construction

There were a total of 12 utterances which showed an afterthought construction. This construction involves two full nominal constituents with the same referent. The afterthought constituent is right adjoined to the clause while the co-referential phrase occurs within the clause. The afterthought is both syntactically complete and intonationally distinct from the main clause; it occurs after a considerable pause and displays its own intonation pattern.

Both constituents in this construction are independent syntactic units however due to what constitutes an NP in this language some examples appear as a single discontinuous noun phrase; consider the examples in (15).

(15)   a.   [plawa-w]   n\mk/u   Yongkulsur-ni
                white-SG.ERG   SG>1NSG.RMT.DUR\gather   Yanggandur-LOC
                ngkei   [belanda-w]
                PROX.LOC   Dutch-SG.ERG

                'The whites assembled (us) here in Yanggandur. . . the Dutch.'

                [004-01 022]

---

[5] Many people do not know their exact year of birth and so providing precise ages is not always possible.

b. mpongke [poi]   n\onta/en   [seki   nel-ni]
    IGN.ALL DIST.LOC 1NSG.RCT.IPFV\sit/ tree.type ground-LOC

'…to that place, we were sitting there…under a seki tree.' [001-01 002]

In (15a) there is an adjective in clause initial position and what appears to be the head noun in word final position with both elements marked with ergative case and sharing the same referent. Similarly in (15b) there is the demonstrative occurring before the verb while what appears to be its NP complement occurs clause finally. In both examples, all the elements which are discontinuous from each other are grammatical as individual constituents and as we saw all word orders are possible. If we were to omit either constituent from each example, this would not negate the grammaticality of these sentences. We have seen that adjectives may serve as the head of the noun phrase in Section 2. Thus it is perfectly possible to have a case marked adjective as a full NP as in (15a). Similarly, we have seen demonstratives occurring without complement NPs (15b).

In fact, the examples in (15) are a minority in the sample. All other examples from the corpus involve elements occurring in both constituents that would be ungrammatical if they were part of a single phrase. The most common examples involve a second head noun (16). (16a) shows two head nouns which have similar semantics although the second one is more precise given the context. Similarly, in (16b) we have the same head *klawo* 'child' being repeated in both the in-situ NP and the afterthought NP. The afterthought NP involves modifier clarifying the gender of the child. If we treated either of these as a single contiguous constituent, the sentence would be ungrammatical. As with above, there would be nothing ungrammatical about omitting either constituent.

(16) a. [mantri] onto k\ole/ngk   [ncokta]
    orderly   can   SG.PST.IRR\exist/ doctor

'So there could've been an orderly…a doctor'   [012-02 028]

b. [klawo nsone]   t\omot/u   piengku, [soro
    child   1SG.POSS SG>3F.RMT.PFV\marry/ 3SG.ERG female
    klawo]
    child

    'He married my child. . . my daughter.'   [010-01 020]

Similarly, example (17a) contains both a possessive pronoun and a case marked
possessor. In (17b), we see an example with a repeated demonstrative. Again, in all
these examples either constituent would be sufficient alone and would be considered
ungrammatical if they were a single contiguous constituent.

(17)   a. [ncuene]   ponto tepi ye   [knume-ne   soi]
            1NSG.POSS truly   just 3.PRS.be Kanum-POSS custom

            'Ours are truly just so. . . the Kanum customs.'   [011-01 038]

        b. [pnm]   mpiya   moro t\orti/u   [pnm
            DIST.INS something how   SG.MID.RMT.PFV\finish/ DIST.INS
            mens-nm]
            fire.INS

            'The thing was finished with it. . . the fire'   [008-01 018]

Not only are both nominal constituents syntactically distinct from the each
other, the afterthought constituent is intonationally distinct as well. All 12 exam-
ples of the afterthought construction show a distinct pause between the clause and
the afterthought of at least 0.4 seconds. Furthermore, the afterthought constituent
has its own intonation pattern signalled by the resetting of pitch to the starting level.
The pitch patterns of the afterthought resembles the same pitch contour of the ut-
terance. Figure 2 is a waveform and pitch analysis of the utterance from (17a). In
the waveform, the top half of the figure, the distinct pause between the main utter-
ance is clearly visible. The bottom half is a pitch analysis conducted in Praat (pitch
range of 11-150Hz). The boundary pitch drop at the right edge of the main clause is
clearly visible. Following the pause, the pitch is reset to the equivilent to the start-

ing pitch, i.e. the left edge of both contours. These same features are perceivable in all examples identified as afterthought constructions.
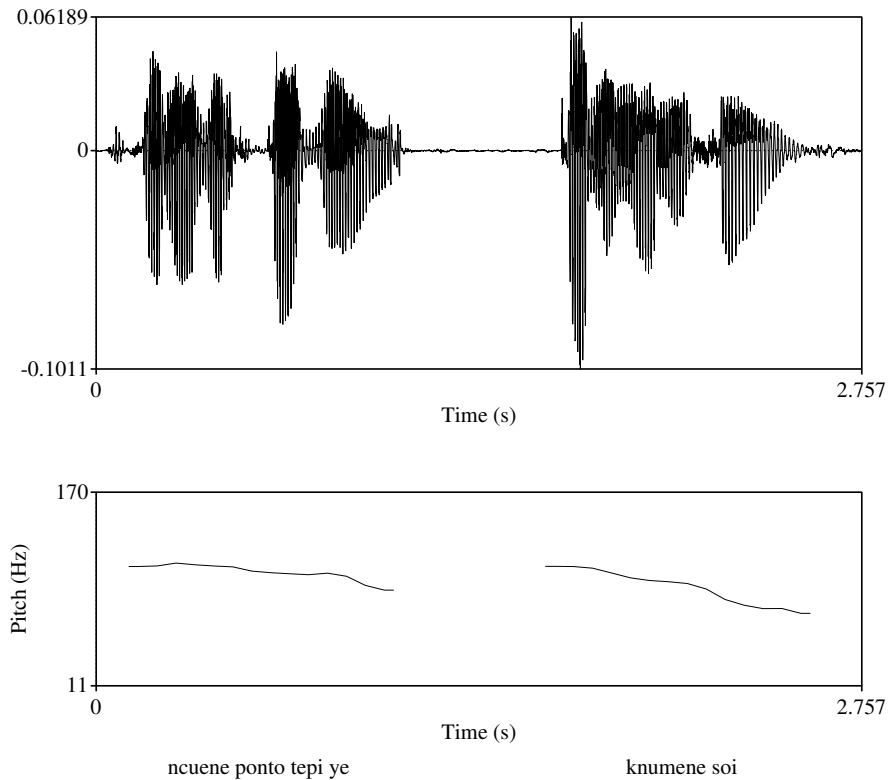

Figure 2: Waveform and pitch analysis of afterthought exemplar.

Common to afterthought constructions cross-linguistically, afterthoughts are typically added to a sentence after it has finished in order to clarify, disambiguate or elaborate (Averintseva-Klisch 2008) (Schultze-Berndt & Simard 2012: 1026) . We see in each example in this section that the afterthought constituent always has a more specific semantics to that of the main clause element. The semantic relation can be hyponymic as in the examples with repeated nouns, e.g. 'white people' and 'the Dutch' (15a), 'orderly'[6] and 'doctor' (16a), 'child' and 'daughter' (16b). In

---

[6] *Mantri* is an Indonesian loan used to refer to whoever is manning the local clinic, be they an unqualified medic, nurse or doctor.

other examples it involves an anaphoric elements in the main clause whose referent it clarified in the afterthought, e.g. 'that place' and 'the seki tree' (15b), 'ours' and 'Kanum customs' (17a), 'it' and 'the fire' (17b).

The afterthought construction, despite giving rise to structures which appear discontinuous (15), is not a true discontinuous noun phrase in the sense that constituent elements would form a single NP if contiguous. Given that the construction allows for the repetition of elements which would be ungrammatical if part of a single noun phrase, it shows us that this construction cannot be interpreted as involving a single noun phrase at some level of representation. Futhermore, the presence of a boundary tone, pause and intonational reset suggests that this afterthought element is not part of the same sentence as the preceding nominal constituent. This same argument has also been argued for afterthought constructions in Jaminjung (Schultze-Berndt & Simard 2012).

3.2 Split DPs

The remaining 24 examples of non-contiguous nominals involve discontinuity between the noun phrase and its demonstrative as in (18).

(18)   [pr]   ngkai   nmaeit s\ow/i   [pi]
        tree.ABS 1SG.ERG already SG>3.HOD.PFV\see DIST.ABS

        'I had already seen that tree.'   [015-01 017]

In this example, we have the head noun, *pr* 'tree', occurring in clause initial position before the subject. A corresponding demonstrative occurs clause finally following the verb. The demonstrative shares the same referent as the noun phrase; it clearly cannot be referring to the subject which is a first person pronoun.

In this section, we will see evidence that the structure we see in (18) is the result of the interaction between the tendency to position topics at the left edge of the clause occurring simultaneously with a demonstrative enclitic which refers

to the topic and occurs following the verb in a reduced phonetic status. Thus, we see what appears to be discontinuity between the NP and demonstrative; however, the internal structure of the NP is always maintained and this is restricted to given information.

3.2.1 The topic clitic

All non-afterthought examples of discontinuity in the corpus involve the demonstrative occurring clause finally. Typically, this is immediately following the verb but also potentially following any other topical demonstratives following the verb (19).

(19)   [kotip] y\owngk/ai   poi   [pi]
        fish   NSG>3.RMT.DUR\place DIST.LOC DIST.ABS

        'We put the fish there.'   [008-01 022]

The demonstrative has a reduced phonetic status. There is no pause between it and the intonation is what we would expect from a basic clause; this can be seen in the waveform and pitch analysis in Figure 3 for the example in (20).

(20)   [nongkai pr] nmaeit y\makr/u   [pi]
        food   tree already SG>3.RMT\burn/ DIST.ABS

        'He had already cooked those foodstuffs.'   [017-01 020]

The demonstrative involves a reduced articulation. The stop /p/ in the distal demonstrative, *pi*, is often lenited to either a fricative or approximate [Φ] or [w]. We see both interspeaker and intraspeaker variation in this. Some speakers are more likely to lenite the stop than others however it also appears to be dictated by speech rate in which faster speech rates involve highly likelihood of lenition. This lenition never occurs with the proximal demonstrative *ngki*. Further investigation into what drives the variation is still needed.
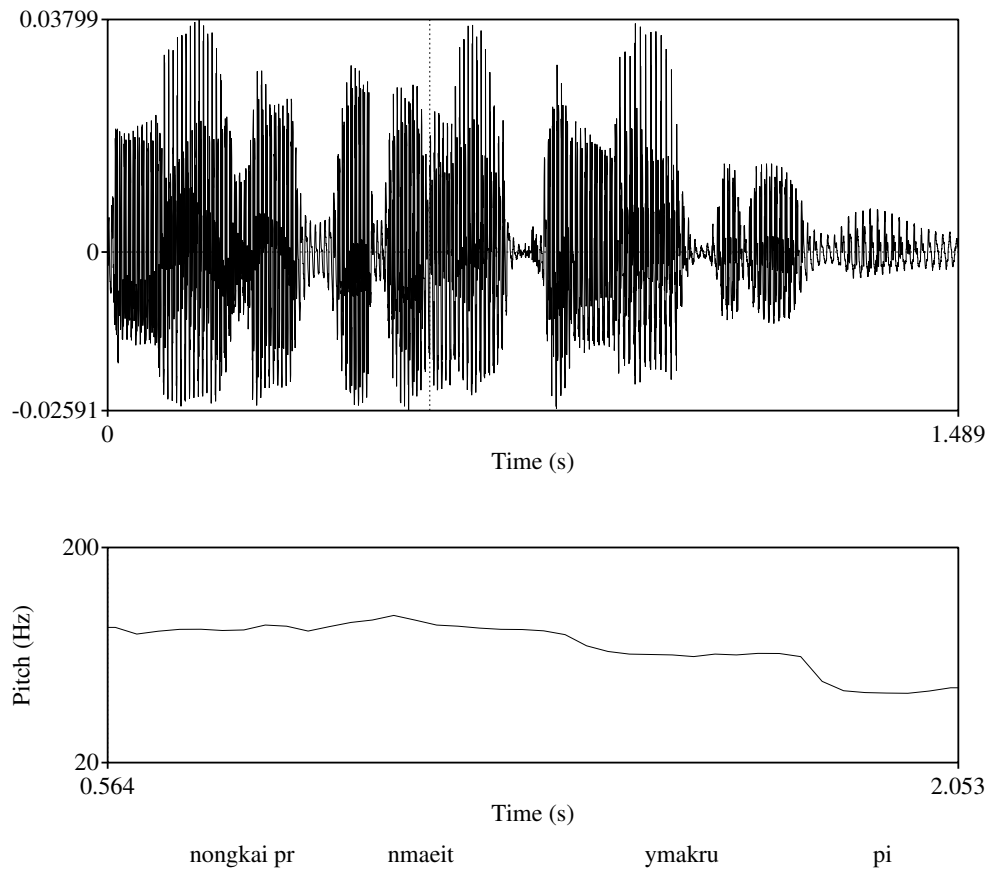
Figure 3: Waveform and pitch analysis of topic clitic exemplar.

We also find encliticised demonstratives occurring without the corresponding NP. Here we have two examples, one with the distal demonstrative (21a) and one with the proximal (21b). (21a) is referring to the old village pastor who is introduced in the previous clause. Example (21b) was produced in response to watching a video of a mourning ritual in which the speaker is describing the arrival of the widow who was also introduced in the previous clause and can be seen by both speaker and addressee on the screen hence the proximal demonstrative is used.

(21)   a.   nmaei pno   belanda-mpa   su\rar/rnt   pi
               before that.time netherlands-ABL SG.RMT.IPFV\be.PL DIST.ABS

               'At that early time, he was from the Netherlands.'   [012-02 017]

   b.   yentnto ponto anni   ngki
               continue truly SG.HOD.come PROX.ABS

               'This one continued coming.'   [013-01 003]

Rather interestingly, the demonstrative clitic never appears with a full DP; only ever an NP. When presented with examples of a demonstrative clitic occurring with a DP in initial position speakers reject it as ungrammatical (22).

(22)   *[[pr]   pi]   ngkai   nmaeit s\ow/i   [pi]
           tree.ABS DIST.ABS 1SG.ERG already SG>3.HOD.PFV\see DIST.ABS

           'I had already seen that tree.'   [015-01 017]

All examples of this construction involve given information which we can describe as discourse topics (Erteschik-Shir 2007). All referents were highly-salient and available referents previously introduced in the text.

To illustrate, let us consider the example in (23). This is a sample of a procedural text in which the speaker describes the correct way to open up a garden site in the forest including the correct order of all ritual feasts which also occur. He starts by introducing the topic of the feast (23a) which is conducted after the forest is first cleared. The speaker then turns to the discussion of preparing the feast, referring to the making of the fire-pit where the food will be cooked. The enclicitised demonstrative is then used to refer once again to the feast (23c) before the speaker moves to discussing measuring the land for dividing the workload (23d). This particular feast is not discussed again in this text.

(23)   a.   sraku   moro kr\ayel/ei,   nmaei **kai** kr\ey/ei
               clearing BG   1NSG.IRR\work.PL, olden feast 1NSG.IRR\work

               'When we make the clearings, we make a feast.'

b. **mens-mens** kr\ey/ei   poi
    fire-fire   1NSG.IRR\work DIST.LOC

    'We make the fire pit there.'

c. [**kai**] kr\ey/ei   [pi]
    feast 1NSG.IRR\work DIST.ABS

    'We make that feast.'

d. sraku   tepi ye pno,   sr\trunant/ei   pngk
    clearing just is that.time, 1NSG>3.IRR\measure DIST.ALL

    'Once its cleared, we measure it up.'   [014-01 008-011]

3.2.2 The noun phrase

All examples of nominals separated from their demonstrative maintain the internal
structure for their NP as established in Section 2. Examples (24) & (25) shows a
modifying noun followed by a head noun. Example (26) shows a basic adjective
modifying a head noun and example (27) shows a derived adjective modifying a
head noun. There are no examples in the corpus of elements within the NP display-
ing alternate orderings.

(24)   [nongkai pr] nmaeit y\makr/u   [pi]
        food   tree already SG>3.RMT\burn/ DIST.ABS

        'He had already cooked those foodstuffs.'   [017-01 020]

(25)   [teya   mpl] opo s\me/i   [pi]
        banana trunk just SG>3.HOD.PFV\give DIST.ABS

        'He had just given him the banana trunk.'   [013-01 109]

(26)   [nmaei brar]   sw\rar/rnt   poi   [pi]
        olden   garden 3.RMT.IPFV\be DIST.LOC DIST.ABS

        'That old garden was there.'   [017-01 020]

(27)   mpyae-t   [prae-wa   were-nm]   sr\makr/nt   [pnm]
       thing-PURP   heat-ADJ   light-INS   SG>3.IRR\burn   DIST.INS

       'For that reason, he'll burn him with the hot sunlight.'   [021-01 011]

All examples of this construction involve the NP element as the first argument in the clause. This occurs preceding any other core constituent and following any sentential adverbial elements (27). As we saw in Section 1.1, word order in Ngkolmpu is syntactically free and so this is not unexpected for topics.

Cross-linguistically, we find a tendency for given information to occur before new (Halliday 1967; Prince 1981). A situation we find confirmed in Ngkolmpu. As we saw, the default word order of Ngkolmpu is AOV; yet all examples from the corpus of non-A arguments occurring before the A argument involve an argument that was previously introduced in the story or immediately available to both speaker and listener. Thus, we see a tendency for topics, when overtly expressed, to occur before other arguments. We can see this in (28a), where we have an O argument, *bpe* 'club', preposed before the A argument. We can see that this is available information as it is introduced earlier in the story (28b).

(28)   a.   bpe   piengku   moro   s\ie/ngk
             club   3.SG.ERG   BG   sg>3.rmt\lift

             'He lifted up the club'   [017-01 040]

        b.   bpe   poi   su\rar/rnt
             club   DIST.LOC   SG.RMT.IPFV\be.PL

             'A club was there.'   [017-01 028]

Thus, when we put together both the tendency to front given information along with marking topics with a demonstrative clitic, we can easily explain such examples as (18) in which the O arguments comes first and its demonstrative final in the clause.

It is worth noting at this point some interesting gaps in the corpus. The first is that all but one example of both a demonstrative clitic and a full NP involve abso-

lutive arguments; the other being in the instrumental case. This is rather interesting as these are not arguments which Donohue claims may undergo scrambling. This point is returned to in Section 4. The other gap is that all examples involve only the distal demonstrative. In some ways, this gap in the data corpus is unsurprising from a statistical perspective. Given the ability of highly available information to be re-coverable context, it seems unlikely for something which is both given information and proximal to both speaker and addressee to be referred with both a full NP and demonstrative.

4   Reconciling Donohue (2011)

The corpus study of the previous section revealed various constructions which po-tentially involve discontinuous or apparently discontinuous structures. In this sec-tion, we return the to Donohue's examples in light of this corpus study. For the sake of brevity, we will focus entirely on the set of examples presented in (14), repeated here in (29).

(29)   a.   [[ntop yrye-w]$_{NP}$ piengku]$_{DP}$ sr\eyerk/nt
              big   man-SG.ERG DIST.ERG   SG>3.FUT.DUR\stalk/

       b.   [ntop yrye-w]$_{NP}$ sr\eyerk/nt   [piengku]$_{DP}$
              big   man-SG.ERG SG>3.FUT.DUR\stalk/   DIST.ERG

       c.   [[yrye-w]$_{NP_i}$ piengku]$_{DP}$ sr\eyerk/nt
              man-SG.ERG DIST.ERG   SG>3.FUT.DUR\stalk/
            [ntop-w]$_{NP_i}$
              big-SG.ERG

       d.   [[ntop-w]$_{NP_i}$ piengku]$_{DP}$ sr\eyerk/nt
              big-SG.ERG DIST.ERG   SG>3.FUT.DUR\stalk/
            [yrye-w]$_{NP_i}$
              man-SG.ERG

       e.   [ntop-w]$_{NP_i}$ mo   [piengku]$_{DP}$ sr\eyerk/nt
              big-SG.ERG wallaby   DIST.ERG   SG>3.FUT.DUR\stalk/
            [yrye-w]$_{NP_i}$
              man-SG.ERG

'That big man will stalk it/a wallaby.'

Example (29a) is the default word order and requires no comment. Example (29b) involves a case marked NP occurring in initial position and a demonstrative following the verb. In this case, this could easily be interpreted as involving a demonstrative enclitic on the verb.

From here things are more complicated as examples (29c) to (29e) were unanimously rejected by speakers when presented for grammaticality judgements. Each example was presented to at least four speakers on at least two non-consecutive occasions to each speaker. In each instance, I spoke the sentence and the speaker would be asked to repeat the sentence and then comment if it was a good sentence in Ngkolmpu; every time these sentences were rejected as ungrammatical. Structurally identical sentences were also rejected. However, the analysis presented below involves these sentences as involving afterthought constructions which, as a repair strategy, are not especially amenable to grammaticality judgements. This could explain why they were rejected. Other afterthought constructions are also rejected by speakers when presented.

With that in mind, I propose the following analyses for sentences (29c) to (29e). Example (29c) involves a standard DP followed by a final case marked adjective. This can be said to be an afterthought construction. Remember that single case marked adjectives are grammatical as an NP, as we saw with elided nouns in 2. Thus, this analysis would account for this structure. A similar analysis is possible for (29d); in this example the adjective is the head of the NP and the noun is a second afterthought NP following the clause clarifying the preceding nominal element.

Example (14e) is the only example which is unexplainable given the data presented in Section 3. One possible analysis is that it involves the interaction of a demonstrative clitic, topic fronting and an afterthought. This could present some way of reconciliation between Donohue's data and the corpus study of this paper.

Under this analysis, *ntopw* 'big.ERG' is the NP occurring in topic position at the front of the clause. The noun *yryew* 'man.ERG' is the afterthought which occurs following the clause. The main abnormality is the presence of the demonstrative preceding the verb.

As mentioned in Section 3.2.2, there are no examples of a demonstrative clitic occurring with a separate NP in the ergative. This can be explained probabilistically; as agents and topics are both highly given and thus likely to be omitted under the standard processes for this. As such, it is a low probability that a topical ergative marked agent would occur in an utterance along with a demonstrative marking deixis. While we have no naturalistic data, elicited examples show the ergative demonstrative occurring procliticised to the verb similarly to the encliticisation of other demonstratives. (30) was presented to three speakers who accepted it as a grammatical sentence; it is identical to the problematic sentence except the final noun has been omitted.

(30)   [ntop-w]   mo   [piengku]   sr\eyerk/nt
        big-SG.ERG   wallaby   DIST.ERG   SG>3.FUT.DUR\stalk/

        'That big one will stalk a wallaby.'

If, then, we treat the final noun *yryew* 'man.ERG' as an afterthought, this would explain the grammaticality judgements of (29e) given that afterthoughts are typically rejected by the speakers I worked with. Thus, I propose that (29e) is an an interaction between a demonstrative enclitic and an afterthought.

This still does not reconcile the lack of demonstrative proclicits in the corpus. However in grammaticality judgement contexts, we find that the ergative demonstrative may occur following the verb similar to the absolutive (31a). Although, it does not appear grammatical to have the absolutive demonstrative occur preceding the verb (31b). So if we pursue this analysis, it seems the ergative may precede or follow the verb while other demonstratives may only follow.

(31)   a.   [ntop-w]   mo   sr\eyerk/nt   [piengku]
            big-SG.ERG wallaby SG>3.FUT.DUR\stalk/ DIST.ERG

            'That big one will stalk a wallaby.'

       b.   *[mo]   yrye-w   [pi]   sr\eyerk/nt
            wallaby man-SG.ERG DIST.ABS SG>3.FUT.DUR\stalk/

            'The man will stalk *that wallaby*.'

We have seen in this section, that potentially the data presented in Donohue
(2011) is reconcilable with the corpus study presented in Section 3. We have sup-
plemented the set corpus data with less natural data. Naturally, data collected in this
manner is is given without context and intonation tests are less reliable; however it
seems necessary given the rarity of certain constructions.

5   Conclusion

The data from the naturalistic corpus shows that discontinuous noun phrases in
Ngkolmpu are highly restricted and that some examples which superficially appear
discontinuous are not truly discontinuous. With this in mind, we have returned
to the data Donohue (2011) used to claim that Ngkolmpu displays discontinuous
noun phrases and saw that these are largely compatible with the corpus data sup-
plemented with minimal experimental data.

The corpus study reveals that Ngkolmpu displays two types of constructions that
can result in apparently discontinuous structures. The first is an afterthought, which
is a repair strategy involving two completed nominal elements yet which share
the same referent. The other is the result of two independent but potentially co-
occurring means for marking topichood: topic fronting and a demonstrative clitic
on the verb. Rather compellingly. however, these constructions are quite rare in the
corpus and occurring in just 47 out of 1500 utterances. This rarity is confirmed by
my own observations in the field however this must be considered anecdotal only.

A more general point can be made about methodology. This paper highlights the necessity of providing naturalistic data which is embedded both in its textual / informational context as well as recording natural production to capture prosodic and intonational elements. Data based on naturalistic data paints a more subtle picture; showing that discontinuity in this case is both rare and contingent on information structure and prosody as well as showing syntactic behaviour which may not be accepted in elicitation. This is not to discount that elicited data which are a necessary part of linguistic field work required to access rarer constructions. However, generalisation must be built on naturalistic data so as to not unduly privilege more marginal constructions.

Acknowledgements

Abbreviations

>	acting on	IRR	irrealist
1	first person	LOC	locative
2	second person	MID	middle voice
3	third person	MOD	modifier
ABL	ablative	N	noun
ABS	absolutive	NSG	non-singular
ADJ	adjective	NUM	number
BG	backgrounding	PFV	perfective
DAT	dative	PL	plural
DIST	distal	POSS	possessive
DUR	durative	PROX	proximal
ERG	ergative	PRS	present
F	feminine	PST	past
FUT	future	PURP	purposive
HOD	hodiernal past	RCT	recent past
INS	instrumental	RMT	remote past
IPFV	imperfective	SG	singular

References

Averintseva-Klisch, Maria. 2008. To the right of the clause: Right dislocation vs. afterthought. In Cathrine Fabricius-Hansen & Wiebke Ramm (eds.), *'subordination' versus 'coordination' in sentence and text: A cross-linguistic perspective*, Amsterdam: John Benjamins Publishing Company.

Carroll, Matthew J. 2012. Languages of southern New Guinea project (LSNG04). Digital collection managed by PARADISEC. [Open Access]. https://doi.org/10.4225/72/56E979083360E.

Carroll, Matthew J. 2016. *The Ngkolmpu Language with special reference to distributed exponence*: School of Culture, History & Language, College of Asia & the Pacific, The Australian National University dissertation.

Dench, Alan & Nicholas. Evans. 1988. Multiple case-marking in Australian languages. *Australian Journal of Linguistics* 8. 1–47.

Donohue, Mark. 2005. Configurationality in the languages of New Guinea. *Australian Journal of Linguistics* 25(2). 181–218.

Donohue, Mark. 2011. Case and configurationality: scambling or mappying? *Morphology* 21. 499–513.

Erteschik-Shir, Nomi. 2007. *Information structure:the syntax-discourse interface*. Oxford: Oxford University Press.

Evans, Nicholas, Wayan Arka, Matthew Carroll, Yung Jung Choi, Christian Döhler, Volker Gast, Eri Kashima, Emil Mittag, Kyla Quinn, Dineke Schokken, Jeff Siegel, Philip Tama & Charlotte van Tongeren. 2017. The languages of southern New-Guinea. In *The languages and linguistics of the New Guinea area*, Berlin: Mouton de Gruyter.

Halliday, M. A. K. 1967. Notes on transitivity and theme in english: Part 1. *Journal of Linguistics* 3(1). 37–81.

Prince, E.F. 1981. Toward a taxonomy of given - new information. In Peter Cole (ed.), *Radical pragmatics*, 223–255. New York: Academic Press.

Schultze-Berndt, Eva & Candide Simard. 2012. Constraints on noun phrase discontinuity in an australian language: The role of prosody and information structure. *Linguistics* 50(5). 1015–1058.

Address for correspondence

Matthew J Carroll
Centre of Excellence for the Dynamics of Language
H.C. Coombs Building
Australian National University

Canberra, ACT, 0200
Australia

matthew.carroll@anu.edu.au
www.matthewjcarroll.com
https://orcid.org/0000-0002-8419-0539