Chapter 7

Austronesian predication and the emergence of biclausal clefts in Indonesian languages

Daniel Kaufman

Queens College & ELA

Information structure is tied up closely with predication in predicate-initial Philippine-type languages. In these languages, subjects are presupposed and the predicate position operates as a default focus position. The present paper argues that there is no need for biclausal focus constructions in these languages due to the nominal properties of their event-denoting predicates. Non-Philippine-type Austronesian languages develop a stronger noun/verb distinction that I argue ultimately gives rise to biclausal focus constructions. The building blocks of biclausal clefts in Indonesian languages are analyzed as well as the nature of predication in Philippine-type languages. Finally, I discuss a paradox in the syntax of definite predicates in Philippine-type languages. In a canonical predication, the less referential portion (the predicate) precedes the more referential portion (the subject). However, when both portions are definite the relation is reversed such that the more referential portion must be initial. I tie this to animacy effects found in other Austronesian languages in which a referent higher on the animacy scale must linearly precede one that is lower.

1 Introduction

Languages vary widely in their strategies for indicating pragmatic relations such as FOCUS and TOPIC. In the simplest case, a language may employ dedicated morphological markers which combine directly with focused or topical constituents. More common perhaps is the use of dedicated syntactic positions, typically on the left periphery of the clause, which host focused or topicalized constituents. Finally, all languages are thought to express basic information structure via prosodic means, although the actual implementation differs significantly from language to language. Parallel to pragmatic relations such as TOPIC and FOCUS, all languages possess a basic SUBJECT-PREDICATE relation which is partly independent of pragmatics but which also intersects with the phenomena of topic, focus and presupposition. While there has been notable success in the definition and analysis of pragmatic relations over the last several decades, the true nature

of the subject-predicate relation remains one of the most fraught topics in the history of linguistics, with its beginnings in the work of Aristotle and Plato. Indeed, as Davidson (2005: 83) states with regard to predication, "It is a mark of Plato's extraordinary philosophical power that he introduced a problem that remained unresolved for more than two millennia." As could be surmised solely from the disagreement among syntactic analyses of English, a robust *cross-linguistic* definition of predicate and subject remains even more elusive. Copular clauses and cleft structures are of special interest here as notional predicates can occupy the canonical syntactic position of the subject (and vice versa) in these sentence types. In the present work, I am concerned with the interplay between the subject-predicate relation and information structure in Austronesian languages. Specifically, I would like to answer the following three questions:

(i) What is the evidence for biclausal cleft structures in Philippine-type[1] versus Indonesian languages? (§3)

(ii) How and why do biclausal cleft structures come into being outside of Philippine-type languages? (§4)

(iii) What does it mean to be a predicate in Philippine-type languages? (§5)

We can briefly preview the answers put forth below. With regard to (i), I argue that a genuine cleft structure in Indonesian languages emerges from a more symmetric Philippine-type system where true biclausal clefts do not exist. In regard to (ii), I show how a distinction between plain modification and modification by relative clause develops in Indonesian languages and how old functional morphology is recruited to mark relative clauses. Finally, regarding (iii), I argue that the predicate-subject relation in Philippine-type languages is determined by the *relative referentiality* of the two basic parts of a proposition similar to copular clauses in more familiar languages. The more referential half of the predication (i.e. the subject) follows the less referential half (i.e. the predicate). An interesting complication is that when both the predicate and the subject are referential, the part of the predication higher on the referentiality/animacy scale precedes the one lower on the scale. That is, the principle which derives the normal predicate-initial order in Philippine-type languages appears to be reversed in these cases.

Languages that sit on the border of the Philippine-type and non-Philippine-type are especially interesting in regard to information structure. In §4, I examine Balantak as a language that appears to be transitioning from monoclausal to biclausal focus constructions. This in turn sheds light on the development of biclausal constructions in languages that have diverged even further from the Philippine-type, such as Malay/Indonesian.

In §2, I attempt to define all the relevant categories in terms that are as theory-neutral as possible. The relevant notions for our purposes are subject and predicate (§2.1), prag-

---

[1]"Philippine-type languages" refer to a typological grouping rather than a geographic or phylogenetic one. It is used here to refer to Austronesian languages with a historically conservative set of (3 or 4) voices (following Blust 2002). Crucially, in Philippine-type languages, these voices are symmetrical in that a predicate can only bear one voice at a time and the agent argument is not demoted in the non-actor voices, as it would be in a canonical passive.

matic relations such as topic, focus and presupposition (§2.2), and the various types of clefts together with their component parts (§2.3).

2 Defining the terms

2.1 Subject and predicate

Several streams in philosophy of language, semantics and even formal syntax have taken a purely taxonomic approach to the notion of predicate with the goal of seeking a unifying trait in these types. The philosophical literature, in particular, is replete with claims such as "predicates ascribe" and "predicates designate". One of Frege's most important contributions to our current understanding of predicate involved viewing it as an element with *unsaturated* arguments; in his words, "not all parts of a thought can be complete; at least one must be unsaturated or predicative; otherwise they would not hold together" (Frege 1892/1997: 193).[2] For Frege, linguistic elements such as names and definite descriptions could not be classified as predicates as they cannot be naturally thought of as having unsaturated arguments in the way that "runs" has a single unsaturated argument and "punches" has two unsaturated arguments. But the fact that languages routinely place definite descriptions, names and even pronouns in the predicate position of clauses that bear all the morphosyntactic hallmarks of canonical predication poses an immediate empirical challenge to Frege's view of predicates as a natural class of *linguistic* elements.[3] Under the direct or indirect influence of Frege, a large body of work in syntax has treated such sentences as something other than pure predications. This has led, for instance, to a taxonomic tradition in the study of copular sentences (Mikkelsen 2011), in which copular clauses can come in specificational, equational and identificational flavors which largely correlate to the referentiality of the "predicate".

For present purposes, the notions SUBJECT and PREDICATE can be defined following type-theoretic predicate logic. Namely, the subject and predicate are the two constituents that combine to yield a TRUTH VALUE. It is not always a trivial matter to determine what types of strings have a truth value and which do not. In the simple case, we can compare the modification relation in (1) with the predication in (2).

(1) Tagalog
     ang   matangkad na   dalaga
     NOM tall   LNK girl
     'the tall girl'

---

[2]As a reviewer notes, an important aspect of Frege's contribution was to assimilate all types of predication to verbal predication. In Fregean semantics, predicates can be defined simply as categories with unsaturated terms.

[3]As noted by Modrak (1985), among others, philosophers have chiefly attempted to explain metaphysical predication, which only bears an incidental relation to linguistic predication. Patterns of linguistic predication across human language have thus not played a major role in philosophical investigations.

(2) Tagalog
Matangkad ang   dalaga.
tall   NOM girl
'The girl is tall.'

Tagalog speakers understand (1) as having a potential reference in the world but not a truth value, whereas the opposite intuition holds for (2) (which nonetheless contains the referring expression *ang dalaga* 'the girl'). In (1), the entire string consists of a single Determiner Phrase (DP) marked with the nominative case determiner *ang*. In (2), the string contains two major phrases, a predicate, followed by the nominative marked DP.

Himmelmann (1986) takes the key feature of predications to be "challengeability": "A predicative structure always allows for – or even demands – a yes-no reaction" (Himmelmann 1986: 26). This view, correctly, I believe, draws a strong line between predication on the one hand and modification, secondary predication and even subordinate predication on the other hand, a distinction which not all theories abide by. I also agree here with Himmelmann in understanding predicates to be crucially a relational concept rather than an inherent property of certain types of linguistic elements.[4]

A particularly vexed question in Philippine linguistics regards which of the two arguments of a basic transitive clause should be considered the "subject", with all possible answers having been posited by different analysts (including "none of the above", see Schachter 1976). In (3), we see three variations on a patient voice clause and in (4) we see the same kind of variations in an actor voice clause.[5] Following a type-theoretic approach, we can see that there is an important difference between the (b) and (c) sentences. In an out-of-the-blue setting, (3b) and (4c) are judged to have truth values but (3c) and (4b) are not. The latter two sentences are not ungrammatical, but they must depend on the preceding discourse to obtain a truth value. That is, as long as anyone ate the tofu, (3b) will be judged true but (3c) cannot be judged as true or false even if we know that Juan ate something. Similarly, for just anyone to have eaten tofu does not make the actor voice sentence in (4b) true. In order for it to be evaluated as true or false, the preceding discourse has to provide the reference for the elided nominative argument.

---

[4]I prefer though to rely on *the potential for a truth value* rather than the notion of challengeability as the latter cannot cleanly apply to imperatives and interrogatives. While some views on questions and imperatives take them to lack truth values, questions and imperatives seem to me best understood as other (non-assertive) things we do with truth values. Declaratives assert that a proposition is true or false; yes-no questions request the hearer to posit a true or false judgement on the proposition; content questions request a value to make a proposition with a variable true; imperatives demand that the addressee make the proposition true. None of these acts could be executed if the proposition itself had no truth value. Thus, just as *Dog* is not a predication, neither is *Dog?* a possible yes-no question, nor *Dog!* a command. I believe these facts can be unified in any approach that sees speech acts as *operations on truth values* rather than restricting truth values to assertions.

[5]In the non-actor voices, the agent is expressed in the genitive case, treated by some as ergative case, while the argument selected by the voice morphology is expressed in the nominative/absolutive case.

(3) TAGALOG

    a. *K<in>áin-Ø ni Juan ang tokwa.*
       <BEG>eat-PV GEN Juan NOM tofu

       'Juan ate the tofu.'

    b. *K<in>áin-Ø ang tokwa.*
       <BEG>eat-PV NOM tofu

       'The tofu was eaten.'

    c. *%K<in>áin-Ø ni Juan.*
       <BEG>eat-PV GEN Juan

       'Juan ate (it).'

(4) TAGALOG

    a. *K<um>áin ng tokwa si Juan.*
       <AV.BEG>eat GEN tofu NOM Juan

       'Juan ate tofu.'

    b. *%K<um>áin ng tokwa.*
       <AV.BEG>eat GEN tofu

       '(S/he) ate tofu.'

    c. *K<um>áin si Juan.*
       <AV.BEG>eat NOM Juan

       'Juan ate.'

The key generalization then is that a predicate must combine with a nominative/absolutive (in Tagalog, *ang* marked) argument to obtain a truth value. On this basis, we can refer to the *ang* phrase as the subject and the clause-initial phrases in the above examples as the predicate. Precisely parallel facts have been observed for several Polynesian languages, such as East Futunan and Tongan (Dukes 1998; Tchekhoff 1981; Biggs 1974). Dukes (1998) sums up the Tongan situation in a way that describes Philippine-type languages equally well:

> An omitted ergative argument need not presuppose any particular referent in the discourse and may in fact be interpreted existentially in much the same way an omitted agent in an English passive is interpreted. When an absolutive is omitted however, it must be interpreted referentially as a null pronoun picking out some previously mentioned individual. As Biggs puts it, native speakers of these languages consider sentences which are missing an absolutive to be 'incomplete', whereas sentences missing ergatives are not.

Note that this definition of subject is completely independent from the "subject" as identified by classic syntactic diagnostics like those posited by Keenan (1976), e.g. binding relations, raising and control, many of which converge on the more agentive argument of a transitive clause, as in English, and only some of which seem to pick out the *ang*

phrase.[6] The concept of the subject-predicate relation as posited above is inherently symmetrical; the subject and predicate are simply the two (topmost) constituents which are combined to yield a truth value. However, few if any human languages treat these constituents symmetrically. There are clearly distinct positions for subject and predicate in the vast majority of described languages in the world.[7] Certain types of copular clauses, however, are apparently reversible in many languages but with subtly different entailments. Jespersen (1937/1984; 1965) notes the distinct meanings of 'my brother' in the following passage:

> Now, take the two sentences:
>
> > *My brother* was captain of the vessel, and
> > *The captain* of the vessel was my brother.
>
> In the former the words *my brother* are more definite (my only brother, or the brother whom we were just talking about) than in the second (one of my brothers, or leaving the question open whether I have more than one). (Jespersen 1965: :153)

Based on a family of similar observations, Jespersen develops the idea that choice of subject and predicate is based on relative familiarity. We can therefore conceptualize predication as an inherently *symmetrical* relation but one whose syntactic expression is highly sensitive to referentiality. That is, the more referential of the two elements in a predication relation will be mapped to a position which we can, following tradition, refer to as "subject" and the less referential of the two will be mapped to a position we call the "predicate".[8] In English, a subject initial language, there is only one way to make a predication between *Mary* and *a linguist*, that is, by treating *Mary* as the subject and mapping *a linguist* to the predicate, as in (5). We say that Philippine-type languages are "predicate-initial" because the less referential component of the predication relation is clause-initial while the more referential component follows it, as shown in (6). Just as in

---

[6]The *ang* phrase relation of Tagalog which I refer to here as subject has, in fact, been given so many names over the years to distinguish it from the subject of a hierarchical argument structure that it is hard to keep track. Among others, we find, "predicate base", "pivot", "topic", and the neutral *ang* phrase. (See Blust 2002, Ross 2006, Kroeger 2007 and Blust 2015 for good summaries of the terminological and conceptual challenges presented by Austronesian voice and case.) I maintain the term subject here because of the familiarity of the subject-predicate relation, which is specifically relevant here. Moreover, the hierarchical relations between co-arguments of a predicate (e.g. the relation between subjects and objects) is not relevant for our purposes and so we can avoid the usual confusion. These are, however, very different relations that should be kept separate terminologically and theoretically, as for instance in Foley & Van Valin (1984).

[7]Diverging from most generative syntacticians, den Dikken (2006) does argue for a symmetrical view of predication but applies the term very broadly to include phenomena that are generally analyzed as modification. Heycock's (2013) overview of predication in generative syntax shows how far this line of work has diverged from the traditional Aristotelian concept of a predicate as part of a bipartite proposition that yields a truth value.

[8]This is also very much in line with the view of predication developed by Williams (1997: 331), who treats referential NPs as potential predicates: "It is sometimes thought that a predicate cannot be 'referential'. It seems to me though that the best we can say is that the predicate is less "referential" than its subject, and that what we really mean is something having to do with directness of acquaintance." Concomitantly, for Williams (1997: 323), "A phrase is not predicational in any absolute sense, but only in relation to its subject."

English, the relative referentiality of the two parts of the predication determines their position in the clause. Reversing the order, as in (6b), results in ungrammaticality.

(5) English
    a. Mary is a linguist.
    b. *A linguist is Mary.

(6) Tagalog
    a. *Abogado si   Jojo.*
       lawyer   NOM Jojo

       'Jojo is a lawyer.'
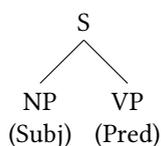
    b. **Si*   *Jojo abogado.*
       NOM Jojo lawyer

I would thus like to take a Jespersonian approach here, which does not rely on purported universal properties of subjects (e.g. "referring") or predicates (e.g. "unsaturated"). Having established this still informal, but hopefully workable, concept of predication, we turn to the impressive flexibility of Philippine languages in mapping lexical categories to the predicate and subject positions, as exemplified in (7). This was noted by Bloomfield (1917) for Tagalog and discussed extensively in the subsequent literature (Gil 1993; Himmelmann 1987; 1991; Foley 2008; Schachter & Otanes 1982; Kaufman 2009a).

(7) Tagalog
    a. *K<um>a~káin   ang laláki.*
       <AV>IMPRF~eat NOM man

       'The man is eating.'

    b. *Laláki ang k<um>a~káin.*
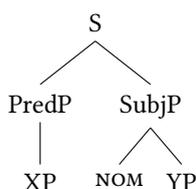   man   NOM <AV>IMPRF~eat

       'The eating one is a man.'

In a very simplistic schema, we can conceive of English and Tagalog differing as in (8) and (9). Whereas the English clausal spine makes crucial reference to lexical categories such as NP and VP, Philippine-type phrase structure seems to refer primarily to the functional categories PredP and SubjP which can in turn host phrases of any lexical category (XP and YP below).[9]

---

[9]Proponents of this view, in one form or another, include Scheerer (1924); Lopez (1937/1977); Capell (1964); Starosta et al. (1982); Egerod (1988); DeWolf (1988); Himmelmann (1991); Lemaréchal (1991); Naylor (1975; 1980; 1995); Kaufman (2009a). Byma (1986), on the other hand, and most subsequent syntacticians (Richards 2000; Rackowski 2002; Aldridge 2004), have defended (or assumed) an analysis in which Tagalog is also structured like (8). Richards (2009a,b) explicitly argues that all predications in Tagalog are mediated by a verbal element but that this element is null in most copular clauses. For reasons of space, I refer the reader to Kaufman (forthcoming) for a rebuttal of this view.

(8)   Eɴɢʟɪsʜ-ᴛʏᴘᴇ ᴘʜʀᴀsᴇ sᴛʀᴜᴄᴛᴜʀᴇ

```
          S
         / \
       NP   VP
     (Subj) (Pred)
```

(9)   Pʜɪʟɪᴘᴘɪɴᴇ-ᴛʏᴘᴇ ᴘʜʀᴀsᴇ sᴛʀᴜᴄᴛᴜʀᴇ

```
            S
           / \
       PredP   SubjP
         |     / \
         XP  NOM  YP
```

In order to understand how phrases are mapped to the predicate and subject position in Philippine-type languages we must first define the crucial pragmatic concepts that come into play. We turn to this in the following subsection.

2.2 Pragmatic relations

Presupposition and focus are often described in shorthand as old information and new information, respectively. While this evokes the right idea, Lambrecht (1994) argues against such oversimplified definitions. I follow Lambrecht's definitions for the relevant categories, as given below.[10]

(10)   **Presupposition, assertion, focus and topic** (Lambrecht 1994: 52, 213, 131)

  a. ᴘʀᴀɢᴍᴀᴛɪᴄ ᴘʀᴇsᴜᴘᴘosɪᴛɪoɴ: The set of propositions lexico-grammatically evoked in a sentence which the speaker assumes the hearer already knows or is ready to take for granted *at the time the sentence is uttered.*

  b. ᴘʀᴀɢᴍᴀᴛɪᴄ ᴀssᴇʀᴛɪoɴ: The proposition expressed by a sentence which the hearer is expected to know or take for granted *as a result of hearing the sentence uttered.*

  c. ꜰoᴄᴜs: The semantic component of a pragmatically structured proposition whereby the assertion differs from the presupposition.

  d. ᴛoᴘɪᴄ: A referent is interpreted as the topic of a proposition if in a given situation the proposition is construed as being about this referent, i.e. as expressing information which is relevant to and which increases the addressee's knowledge of this referent.

---

[10] Abbott (2000) offers an alternative view of presuppositions as *non-assertions* rather than information known to the hearer. This may well fit the Austronesian data better but I leave this question to further work. A good overview of the issues and literature surrounding presuppositions is found in Kadmon (2000).

While the Tagalog *ang* phrase is often referred to as "topic" in different analytic traditions, it has been shown clearly by Naylor (1975), Kroeger (1993) and Kaufman (2005) to have no inherent pragmatic status beyond its definiteness or referentiality. Tagalog and, it would seem, all other Philippine languages have a bona fide topic position on the left periphery. In Tagalog, the fronted pragmatic topic, is followed either by the topic marker *ay* or a pause. In Tagalog, but not all Philippine languages, there is also a dedicated focus position on the left periphery of the clause which hosts oblique phrases, exemplified in (11).

(11) Tagalog
      [*Sa Manila*]$_{FOC}$=ka=ba *p<um>unta*?
       OBL Manila=2S.NOM=QM <AV.PRF>go

      'Did you go to MANILA?' ('Was it to Manila that you went?') (Kaufman 2005: 182)

The presence of an oblique phrase in the focus position in the left periphery bifurcates the sentence into a focus and a presupposition. In the above, 'to Manila' is the focus and it is presupposed that the addressee had gone somewhere. The question would be inappropriate if this information was not already part of the discourse in the same way that 'Was it to Manila that you went?' would be inappropriate in an out-of-the-blue context. On the other hand, a phrase in the left peripheral topic position followed either by the topic marker or a pause, needs to either be in the discourse already or contrasted with a similar argument that belongs to the same set. In (12), the fronted oblique phrase can serve as a contrastive topic, pragmatically akin to English, 'What about Naga, have you gone there?'. Note that the topic is further outside the clause and thus does not host second position clitics. Furthermore, it does not trigger a presupposition. The question in (12) is still felicitous without it being in the discourse that the addressee went somewhere.

(12) Tagalog
      [*Sa Naga*], *p<um>unta=ka=ba*?
       OBL Naga   <AV.PRF>go=2S.NOM=QM

      'To Naga, did you go (there) already?'

With this brief introduction we are now prepared to turn to the mapping of these pragmatic categories on to phrase structure across several Austronesian languages.

2.3 Clefts

Cleft structures make use of the subject-predicate relation to satisfy requirements of information structure. Subjects are canonically (but not necessarily) topic-like and predicates canonically (but not necessarily) align with the focused constituent of a clause. Thus, focused subjects and presupposed predicates constitute non-canonical alignments which languages may either tolerate or avoid by means of various syntactic mechanisms. Languages with a high tolerance for focused subjects, such as English, will allow such a subject to be merely marked by intonation, as in (13a). However, an option also exists to

shift such a focused subject into the predicate position, as in the *it*-cleft sentence in (13b) (Lambrecht 1994).

(13)   a.   Only [John]$_{FOC}$ knows Jane.
       b.   It's only [John]$_{FOC}$ who knows Jane.

At the same time, the presupposition of an English *it*-cleft is demarcated syntactically by means of a relative clause. Thus, while both (13a) and (13b) contain a presupposition 'X knows Jane', its pragmatic status is only reflected syntactically in (13b), by means of the relative clause *who knows Jane*. The English *it*-cleft can thus be said to create a more transparent mapping between the syntactic and pragmatic structure of the clause.

Other languages do not tolerate non-canonical mappings such as that in (13a). For instance, the Malay/Indonesian adverb *saja* 'only', which must combine with a focused constituent preceding it, cannot associate with a subject in a simple declarative clause, as shown in (14a). Instead, a cleft structure is required in which the presupposed predicate is packaged as a relative clause, as shown in (14b).

(14)   Indonesian
       a.   *Presiden*   (*saja) *bisa menilai*   *kinerja menteri.*
            president   only can AV:evaluate output minister

            'A president can evaluate a minister's output.'

       b.   *Presiden   saja   yang bisa menilai   kinerja menteri.*
            president only RELT can AV:evaluate output minister

            'Only a president can evaluate a minister's output.'

As seen in (15), no special manipulation is required in order to narrowly focus the predicate or a part thereof, as this respects the canonical mapping between predicate and focus.

(15)   Indonesian
       Presiden   bisa menilai   kinerja menteri   saja.
       president can AV:evaluate output minister only

       'A president can only evaluate a minister's output.'

Most interestingly, we find that the Austronesian tendency to express presuppositions syntactically manifests itself in Philippine English, as well. Whereas English can employ prosodic focus alone in a sentence like (16), Philippine English will invariably employ a TH-cleft (to be introduced below) in the same function, as seen in (16) and (17). The Tagalog equivalent is given in (18).

(16)   US English
       John will carry your bag.

(17)   Philippine English
       John will be the one to carry your bag.

(18) Tagalog
*Si Juan **ang** mag-da~dala   ng   bag mo.*
NOM Juan NOM AV-IMPRF~carry GEN bag 2SG.GEN
'JUAN will carry your bag.' (Lit. 'Juan will be the one to carry your bag.')

Clefts thus function to transparently bifurcate the sentence into a focus and a presupposition. As seen in the difference between English and Indonesian above, languages differ as to the extent to which they require such transparency. In terms of the syntactic hallmarks of cleft sentences, Lambrecht (2001) offers the following definition:

(19) CLEFT CONSTRUCTION (Lambrecht 2001: 467)
A cleft construction is a complex sentence structure consisting of a matrix clause headed by a copula and a relative or relative-like clause whose relativized argument is coindexed with the predicative argument of the copula. Taken together, the matrix and the relative express a logically simple proposition, which can also be expressed in the form of a single clause without a change in truth conditions.

There are two components in (19). The structural component defines clefts as a biclausal structure containing a relative clause in a larger copular sentence. The semantic component of the definition relates clefts to simpler monoclausal sentences by virtue of their similar meaning. The cleft and the monoclausal structure differ in information structure and implicature but not in their basic truth conditions.

Lambrecht advocates for a taxonomy of English clefts as in (20), where caps indicates focus:

(20) CLEFT TYPES (Lambrecht 2001: 467)
a.	I like CHAMPAGNE.	Canonical sentence
b.	It is CHAMPAGNE (that) I like.	IT cleft
c.	What I like is CHAMPAGNE.	WH cleft (pseudo-cleft)
d.	CHAMPAGNE is what I like.	Reverse WH cleft (reverse/inverted pseudo-cleft)

In (20a) we find the canonical monoclausal sentence which is roughly equivalent in its truth conditions to the following clefts. The *it*-cleft places the focused phrase in the predicate position of a copular clause in which a dummy pronoun is the subject. The presupposition is packaged as a relative clause.[11] The types in (20c) and (20d) are termed alternatively *wh*-clefts or pseudo-clefts (although I will use only the latter term in the following). Here, there is no dummy subject. A relative clause headed by an interrogative pronoun like 'what' is in a copular construction with a focused phrase. In English, this type is reversible so that the presupposition can be the subject of the matrix clause, as in the standard pseudo-cleft exemplified in (20c), or the predicate of the matrix clause, as

---

[11]The precise relation between the two clauses in the English *it*-cleft has been subject to rather intense scrutiny, summed up recently by Reeve (2010).

in the inverse pseudo-cleft exemplified in (20d). To Lambrecht's taxonomy, we can add the "TH-cleft", in (21), first identified as a separate type by Ball (1977). Here, the relative clause modifies a definitely determined semantically bleached noun phrase, e.g. 'the one', 'the thing', etc.

(21)   a.   The one/thing I like is CHAMPAGNE.   TH-cleft
       b.   CHAMPAGNE is the one/thing I like.   Reverse/inverse TH-cleft

The syntax of clefts accommodates information structure in several ways. In structures like the English *it*-cleft (*It's JOHN who bit me*), part of the focus semantics derives from mapping a phrase to the object position of a copular structure, a more hospitable position for focused material than the subject position. In all types of cleft sentences, the presupposition is clearly demarcated by packaging it as a relative clause of some type.

   It is clearly not the case that relative clauses always contain a presupposition outside of cleft sentences. For instance, the relative clause subject in (22) does not presuppose that someone will come after closing time. The sentence in (23), however, does entail such a presupposition, and this shows that it is the determiner or demonstrative that gives rise to the presupposition rather than anything inherent to the relative clause itself.[12]

(22)   Any customer who arrives after closing time will not be served.

(23)   I will take care of those customers who arrive after closing time.

In addition to determiners of a relativized noun, a phrase headed by an interrogative pronoun can be at least partly responsible for projecting a presupposition. A certain chess hustler in Greenwich Village used to rile his opponents with the following rhetorical question during the heat of a match:

(24)   Do you know what I like about your game? Absolutely nothing!

The jarring quality of the answer is the effect of canceling the presupposition in the question. The infelicity of the following cleft sentences in (25) makes this clear. That the presupposition does not come directly from the relative structure can again be seen in the felicitous (26), which contains a relative clause headed by the quantifier 'nothing', and which carries no presupposition.

(25)   a.   %Nothing is what I like about your game.
       b.   %What I like about your game is nothing.

(26)   There is nothing that I like about your game.

---

[12] See Kroeger (2009) for a similar point with regard to mistaken assumptions about headless relative clauses. As Kroeger shows for Tagalog, the presuppositions in such constructions are triggered by the determiners rather than the relative structure itself. Note that Kroeger (1993) claims that the subject actually precedes the predicate in Tagalog translational equivalents to English cleft sentences. On this view, *sino* 'who' would be the subject in a sentence like (27). Kroeger (2009: fn.3), however, recants this position and views (apparent) headless relatives like *ang dumating* 'the one who arrived' in (27) as being in subject position. This change in perspective was prompted by Tagalog's similar behavior to Malagasy as analyzed by Paul (2008) and Potsdam (2009), who offer several pieces of evidence for the predicatehood of (non-adjunct) interrogative phrases.

We can say then that relative clauses pave the road for presuppositions without necessarily triggering them directly. As we will see in the following section, what triggers the presuppositional reading in putative Philippine-type clefts is the definite semantics of the nominative case marking determiner (e.g. Tagalog *ang*). Unlike English, interrogative pronouns are never employed for this purpose in Philippine-type languages.

3 The syntactic structure of Austronesian clefts

A key point of variation between Philippine-type and non-Philippine-type Austronesian languages can be exemplified with the following example form Tagalog (27) and formal Indonesian/Malay (28).

(27) Tagalog

    a. *Sino **ang** d<um>ating?*
        who NOM <AV.BEG>arrive

        'Who arrived?'

    b. *D<um>ating   **ang** guro.*
        <AV.BEG>arrive NOM teacher

        'The teacher arrived.'

(28) Formal Indonesian/Malay

    a. *Siapa **yang** datang?*
        who RELT arrive

        'Who arrived?'

    b. *Datang abang-nya...*
        arrive   elder.brother-3S.GEN

        'His brother arrived...' (Hikayat Pahang 128:9)

Nearly all Philippine-type languages require some form of case marking on clausal arguments. In (27), this can be seen for Tagalog in the case marking determiner *ang*, which can be glossed nominative or absolutive (see Kaufman 2017 for discussion), but whose function is also tightly bound up with a definite/specific reading of the following NP (Himmelmann 1997).[13] There are two crucial things to note in this comparison. First, while Philippine-type languages require such case markers, only few Austronesian languages of Indonesia employ case marking on arguments (see Himmelmann 2005). The relativizer *yang* is strongly preferred in the subject question in (28a) but would not be acceptable in (28b) and can thus be easily distinguished from *ang* in Tagalog and the equivalents in other Philippine languages. Second, the case markers of Philippine languages do not discriminate between apparent verbal and nominal complements. Constantino (1965) was

---

[13]The precise semantics of *ang* has been debated in the literature. A non-definite reading can be obtained in Tagalog with *ang isang...* NOM one:LNK. However, without the presence of the numeral *isa* 'one', felicitous use of *ang* requires familiarity on the part of the hearer. For this reason, I maintain that definiteness, rather than specificity, best captures the pragmatic function of *ang*.

Table 1: Philippine sentence patterns following Constantino (1965)

Tagalog	kina:ʔin	naŋ	ba:taʔ	**aŋ**	maŋga
Bikolano	kinakan	kan	a:kiʔ	**aŋ**	maŋga
Cebuano	ginka:ʔun	han	bataʔ	**aŋ**	maŋga
Hiligaynon	kinaʔun	saŋ	ba:ta	**aŋ**	pahuʔ
Tausug	kyaʔun	sin	bataʔ	**in**	mampallam
Ilokano	kinnan	dyay	ubiŋ	**ti**	maŋga
Ibanag	kinan	na	abbiŋ	**ik**	maŋga
Pangasinan	kina	=y	ugaw	**su**	maŋga
Kapampangan	pe:ŋa=na	niŋ	anak	**iŋ**	maŋga
	eat:PV.PRF	GEN	child	NOM	mango

'The child ate the mango.'

Table 2: Philippine sentence patterns following Constantino (1965)

Tagalog	maŋga	**aŋ**	kina:ʔin	naŋ	ba:taʔ
Bikolano	maŋga	**aŋ**	kinakan	kan	a:kiʔ
Cebuano	maŋga	**aŋ**	ginka:ʔun	han	bataʔ
Hiligaynon	pahuʔ	**aŋ**	kinaʔun	saŋ	ba:ta
Tausug	mampallam	**in**	kyaʔun	sin	bataʔ
Ilokano	maŋga	**ti**	kinnan	dyay	ubiŋ
Ibanag	maŋga	**ik**	kinan	na	abbiŋ
Pangasinan	maŋga	**su**	kina	=y	ugaw
Kapampangan	maŋga	**iŋ**	pe:ŋa=na	niŋ	anak
	mango	NOM	eat:PV.PRF	GEN	child

'It was the mango that the child ate.'

the first to show that this is a far reaching characteristic of Philippine languages with the comparisons in Table 1 and Table 2. In no Philippine language do putative pseudo-clefts contain an overt relative marker, a *wh*-element, a dummy head noun, or any extra sign of nominalization. The voice marked words that serve as predicates in Table 1 are simply bare complements to the determiner in Table 2.

It turns out there is good reason for this symmetry. Starosta et al. (1982) first noted that the well-known Austronesian "voice" paradigm appeared to involve a reanalysis of nominalizations as voice markers, as shown in (29). [14]

---

[14]Unlike the rest of the forms in (29), the actor voice morpheme *<um> is not thought to have developed from a nominalizer, as it can be reconstructed to the proto-Austronesian verbal paradigm (Ross 2002; 2009; 2015).

(29)   Austronesian voice morphology
   -en PATIENT NOMINALIZER > PATIENT VOICE
   -an LOCATIVE NOMINALIZER > LOCATIVE VOICE
   Si- INSTRUMENTAL NOMINALIZER > INSTRUMENTAL VOICE
   <um> AGENT VOICE/NOMINALIZER

I argue in Kaufman (2009a) that the large number of morphological and syntactic similarities between nouns and verbs can be attributed to the reanalysis in (29). This receives further support from Ross (2009), who shows that Puyuma, a Formosan language, maintains a division of labor where the verbs of Philippine-type main clauses are restricted to relative clauses. Another set of verbal morphology, now only used in a subset of Philippine-type languages for imperatives and subjunctives, is used to mark main clause verbal predicates in Puyuma.[15] It seems then that the reanalysis of relative clauses as main clause predicates in an earlier Austronesian proto-language had the effect of erasing any significant differences between relative clauses and main clauses in the daughter languages. If words formed with the morphology in (29) are nominalizations, it is no surprise that they can serve as direct complements of determiners such as seen above in Table 2. There is no need to relativize the verb phrase in sentences such as those in Table 2 if the verb is already akin to a thematic nominalization. To make this concrete, we could compare the patient voice morpheme in (29) to English *-ee* in *employee*. English allows for the two semantically similar sentence in (30).

(30)   a.   George is the one Jane employs.
       b.   George is Jane's employee.

Clearly, direct relativization from a finite clause is far more common and productive in English than thematic nominalization. But in Austronesian, thematic nominalization, as in (30b), was developed to an unusual degree for the purpose of forming relativizations and these then spread to main clauses.[16] A consequence of this, particularly important for focus constructions, is that apparent clefts in Philippine-type languages are monoclausal, just as English (30b) is monoclausal. The key facts are reviewed below.
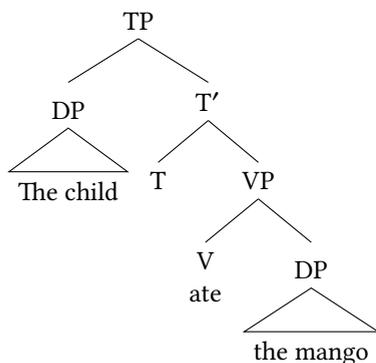
3.1 Apparent Philippine-type clefts: monoclausal or biclausal?

A reasonable analysis of the English pseudo-cleft is shown in (32), which can be compared to the canonical monoclausal declarative sentence in (31).
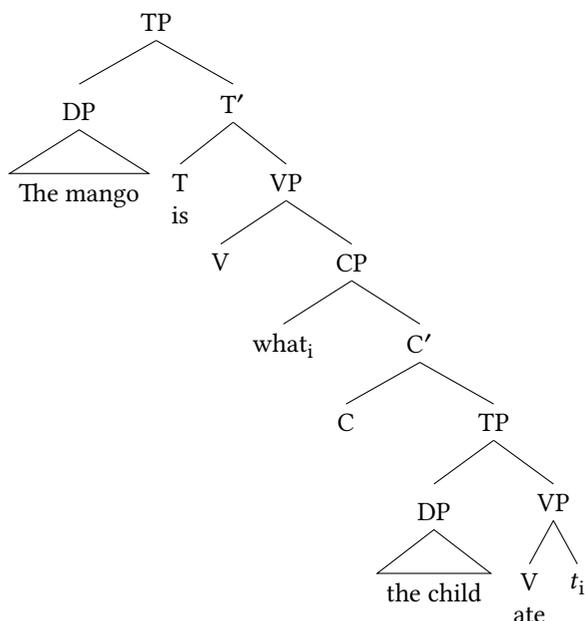
---

[15]Forms which take this set of morphology, referred to as the dependent paradigm by Wolff (1973) and the non-indicative paradigm by Ross (2002), cannot serve as the complement of a determiner or case marker in Philippine languages (Kaufman 2009a: 25).

[16]While this is unusual, it is not unique. Shibatani (2009) notes typological similarities between the use of thematic nominalizations in Austronesian and similar phenomena in Qiang, Yaqui, Turkish and Quechua. With regard to terminology, I refer below to these historically nominalized predicates, i.e. "Philippine-type verbs", as participles rather than verbs or nouns. The term participle is preferable because these forms have an intermediate status between plain nouns and the historical verbs of Austronesian, the latter which have been relegated to non-indicative contexts in Philippine languages.
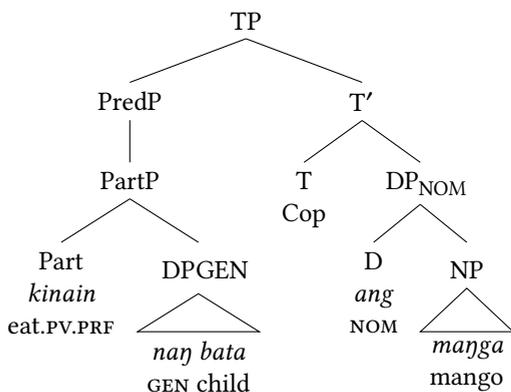
(31)

```
                    TP
              ╱          ╲
           DP              T′
         ╱    ╲          ╱    ╲
     The child      T        VP
                          ╱      ╲
                       V           DP
                      ate        ╱   ╲
                            the mango
```

(32)

```
                    TP
              ╱          ╲
           DP              T′
         ╱    ╲          ╱    ╲
     The mango      T        VP
                    is     ╱      ╲
                         V          CP
                              ╱          ╲
                          whatᵢ            C′
                                        ╱      ╲
                                     C            TP
                                              ╱      ╲
                                           DP          VP
                                         ╱    ╲       ╱  ╲
                                     the child     V     tᵢ
                                                  ate
```
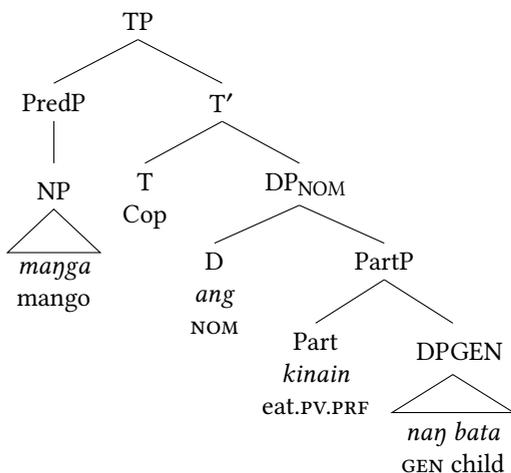
The English pseudo-cleft is considered biclausal because it contains two separate extended projections of a verb phrase, headed by *is* in the matrix clause and *ate* in the relative clause in (32). Each domain can mark categories like tense, negation and agreement independently. In contrast, the monoclausal (31) only contains a single domain for tense, negation and agreement. There is little reason to believe such a distinction exists in Philippine-type languages. The translational equivalents of (31) and (32) in Tagalog both appear monoclausal, as suggested by the analysis in (33) and (34). The only difference is that the participle (descended historically from a nominalization) is in the predicate position in (33) and in the subject position in (34). We can treat both cases, however, as copular clauses, indicated by the (null) Cop in both structures.

(33)

```
                        TP
            ┌───────────┴───────────┐
         PredP                      T′
           │                 ┌───────┴───────┐
        PartP                T            DP_NOM
    ┌─────┴─────┐           Cop         ┌───┴───┐
  Part        DPGEN          D          NP
 kinain      ╱────╲         ang       ╱───╲
eat.PV.PRF  naŋ bata       NOM       maŋga
           GEN child                 mango
```

(34)

```
                     TP
          ┌──────────┴──────────┐
        PredP                    T′
          │              ┌───────┴───────┐
          NP             T            DP_NOM
        ╱───╲           Cop        ┌─────┴─────┐
       maŋga            D         PartP
       mango           ang    ┌─────┴─────┐
                       NOM   Part        DPGEN
                            kinain      ╱────╲
                          eat.PV.PRF   naŋ bata
                                      GEN child
```

One of the few arguments that has been adduced in favor of a biclausal structure for sentences such as (34) is a putatively asymmetric pattern of clitic placement.[17] As discussed

---

[17]The notion that apparent clefts in Philippine languages are biclausal is widespread although often not explicitly argued for. Nagaya (2007: 348), for instance, analyzing Tagalog information structure in an RRG framework, states "A cleft construction in Tagalog is an intransitive clause where its single core argument is a headless relative clause, and its nucleus is a noun phrase coreferential with the gap in the headless relative clause." as illustrated in his (i), where *S* represents a gap in the relative clause.

(i) *Si   Boyet_i ang [p<um>atay [S_i] kay Juan].*
     NOM Boyet NOM <AV>kill   OBL Juan
     'The who killed Juan is Boyet.' (Nagaya 2007: 348)

Even in non-derivational frameworks such as RRG, the gap strategy employed commonly for relative clauses in Indo-European languages is typically applied to the analysis of Tagalog without consideration of alternative analyses.

in detail in Kaufman (2010b), Tagalog pronominal clitics are positioned after the first prosodic word in their syntactic domain. Aldridge (2004: 320), assuming that such clitics strictly take the clause as their domain, presents the data in (35) as an argument for the biclausal structure of apparent clefts. If such sentences were monoclausal, it would stand to reason that clitics could follow the interrogative directly as in (35b), but such a pattern is ungrammatical.

(35) Tagalog (Aldridge 2004: 319)

    a. *Ano   ang   g<in>a~gawa=mo*?
       what NOM <BEG>IMPRF~do.PV=2S.GEN

       'What are you doing?'

    b. **Ano=mo*   *ang   g<in>a~gawa*?
       what=2S.GEN NOM <BEG>IMPRF~do.PV

Second position clitics, however, are not only clause-bound; they are also bound within the DP, as can be seen in the following comparison with the possessive clitic *=ko* 1SG.GEN. With a bare predicate like *kaibigan* 'friend', as in (36a), the possessive clitic attaches to the first element in the clause, in this case, negation. When the predicate is a case marked DP, as in (36b), the associated genitive clitic cannot take second position in the clause and must attach after the first prosodic word within the DP.

(36) Tagalog

    a. *Hindi*[*=ko*]*=siya*   *kaibigan*[?*=ko*].
       NEG=1S.GEN=3S.NOM friend=1S.GEN

       'He is not a friend of mine.'

    b. *Hindi*[**=ko*] *siya*   *ang   kaibigan*[*=ko*].
       NEG=1S.GEN 3S.NOM NOM friend=1S.GEN

       'He is not the friend of mine.'

Similarly, in an event-denoting predication such as (37), second position clitics cannot follow the predicate when they originate within a case-marked DP.

(37) Tagalog
    Na-dapa[**=ko*] *ang   kapatid*[*=ko*].
    BEG-fall=1S.GEN NOM sibling=1S.GEN

    'My sibling fell.'

Aldridge (2004), citing data similar to (36), essentially comes to the same conclusion.[18] But if this generalization is correct, then the earlier clitic argument from (35) for a bi-clausal cleft structure is neutralized. Clitics are unable to escape from a DP and thus the genitive clitic in (35), representing an agent embedded in a nominative phrase, cannot follow the interrogative.

---

[18] Aldridge (2004: 262): "I assume that DP is a strong phase, not permitting movement from it. However, a predicate nominal is not, so the clitic would be able to move."

3.2 True biclausal clefts in Austronesian languages

The nominal properties of "Philippine-type verbs" is largely lost south of the Philippines (Kaufman 2009b). Consequently, Malay, even in its earliest attested stages, does distinguish relative clauses syntactically through the use of *yang*. As shown earlier in (28), Indonesian-type relativizers like *yang* are functionally distinct from Philippine-type case marking determiners. We can further see in (38) and (39) how Indonesian-type relativizers are distinguished syntactically from the "linker" or "ligature" found in most Philippine-type languages. First, *yang* is not required to mediate adjectival modification, as seen in (38a). Second, it can head a phrase without a preceding noun, as seen in (38b). The Philippine linker/ligature differs on both of these counts. It must mediate all instances of modification, as shown in (39a) and cannot surface without a preceding phrase.[19]

(38) Indonesian

    a. *rumah* (*yang*) *besar*
       house   RELT big

       'big house'

    b. (*yang*) *ini*
       RELT this

       'this one'

(39) Tagalog

    a. *bahay* *(*na*) *malaki*
       house   LNK big

       'big house'

    b. (*na) *ito*
       LNK this

       'this'

A relative clause referring to the agent is built on an actor voice VP with the addition of the relativizer *yang*, as in (40a). As can be seen in (40b), the plain VP cannot stand in subject position with the same function.[20]

    The presence of a dedicated relativizer is one crucial piece of evidence for the biclausal nature of the construction. An additional piece of evidence is the optional presence of the copular element *adalah*.

---

[19]See Yap (2011) for a further discussion of *yang* and its expanding functions in the history of Malay.

[20]Verb phrases can also stand in subject position, typically with the help of a demonstrative, when functioning as event nominalizations, as in (i).

    (i)   TP[DP[VP[Menilai   *kinerja mentri*] *itu*] *susah*].
           AV:evaluate output minister that difficult

       'Evaluating the output of ministers is difficult.'

(40) Indonesian

    a. *Yang menilai   kinerja menteri adalah Presiden.*
       RELT AV:evaluate output minister COP   president

    '(The one) who evaluates the output of a minister is the president.'[21]

    b. **Menilai   kinerja menteri adalah Presiden.*
       AV:evaluate output minister COP   president

The innovation of a copula in Indonesian languages has yet to be studied systematically. The copula *adalah* was innovated in the attested history of Malay from a presentative use of the existential *ada* in combination with the emphatic *lah*. Although English-like cleft constructions employing both the copula and a relative clause can be found in modern Indonesian, there remain restrictions on the use of the copula that are not well understood. Specifically, we find that the copula is rejected in questions like (41b), a constructed minimal pair with the attested (41a).

(41) Indonesian

    a. *Dia adalah yang di-tua-kan   di   antara sesamanya.*
       3S   COP   RELT PV-old-APPL PREP among colleague

    'It's him who is treated as an elder among colleagues.'[22]

    b. *Siapa (*adalah) yang di-tua-kan?*
       who   COP   RELT PV-old-APPL

    'Who is treated as an elder?'

In line with the historical development of *adalah*, it is likely that it selects for a focused complement or at least avoids a presupposed one. This is supported by the fact that the copula is again possible when the interrogative is in-situ, as in (42).[23]

(42) Indonesian
    Yang di-tua-kan   (adalah) siapa?
    RELT PV-old-APPL COP   who

    'The one treated as an elder is who?'

The use of a dedicated relativizer and copula in Indonesian (non-adjunct) content questions and focus constructions shows that this language has developed bona fide biclausal cleft sentences where Philippine-type languages still employ an equational monoclausal structure. Unfortunately, the difference between Philippine-type and non-Philippine-type Austronesian languages in this regard has not been given much attention by syntacticians. The default hypothesis has treated languages like Indonesian as simply having

---

[21]http://nasional.republika.co.id/berita/nasional/politik/16/01/06/o0iwuo354-jokowi-yang-menilai-kinerja-menteri-adalah-presiden

[22]http://nasional.kompas.com/read/2014/09/19/06431611/Artidjo.Korupsi.Kanker.yang.Gerogoti.Negara
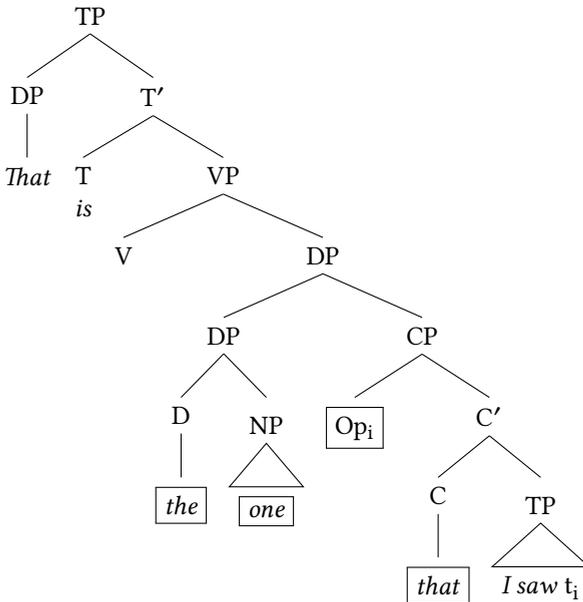
[23]The ungrammaticality of post-interrogative copulas and copula stranding is not addressed by Cole & Hermon (2000).

overt markers for what are null functional elements in Philippine-type languages.[24] In the next section, we explore how the morphological glue of biclausal constructions is recruited from existing lexical and functional elements as part of a larger argument that such constructions are relatively recent innovations in the history of Austronesian.

4 How to jerry-rig an Austronesian biclausal cleft

We can posit a structure such as the one in (43) for an English TH-cleft (introduced earlier in §2.3). The focus here lies in the subject position while the presupposition is packaged as a DP containing a relative clause. Note that there are multiple elements within the presupposition that are special to this construction.

(43)

The DP proper contains a determiner and a semantically bleached noun, in this case *one*. The modifying CP contains a complementizer *that* and, ostensibly, a null operator in the position otherwise reserved for interrogative elements. That these layers are distinct

---

[24]Potsdam (2009) enumerates the pseudo-cleft analyses proposed for wh-questions across a diverse set of Austronesian (including both Philippine-type and non-Philippine-type) languages: Palauan (Georgopoulos 1991), Malay (Cole et al. to appear), Indonesian (Cole et al. 2005), Tsou (Chang 2000), Tagalog (Kroeger 1993; Richards 1998; Aldridge 2004; 2002), Seediq (Aldridge 2004; 2002), Malagasy (Paul 2001; 2000; Potsdam 2006a,b), Maori (Bauer 1991; 1993), Niuean (Seiter 1980), Tuvaluan (Besnier 2000), Tongan (Otsuka 2000; Custis 2004). Chung (2010) specifically traces the analysis of content questions in Philippine-type languages as pseudo-clefts to Seiter (1975). While such analyses appear well supported for many non-Philippine-type languages, it does not seem justified to assume the same structure for Philippine-type languages.

is seen both in historical stages of English, as exemplified in (44), and in non-standard modern English, (45).
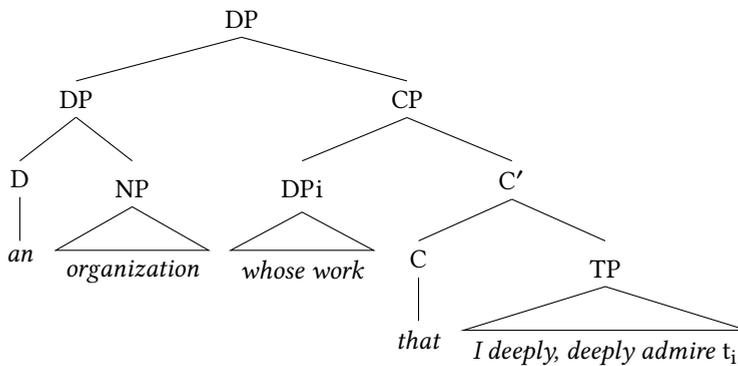
(44) Middle English (Chaucer's *Prolog* 836, cited in Curme 1912)
      He which   that   hath the shortest shall beginne.
      he RELT.PRON COMP has   the shortest shall begin

      'He who has the shortest shall begin.'

(45) Here I am, in this room, because of an organization whose work that I deeply, deeply admire.[25]

In (46), the DP 'an organization', is modified by a CP which contains both an interrogative phase, 'whose work', moved to its left periphery and the complementizer 'that'.

(46)

Given the distinct roles and positions of the determiner, dummy noun, interrogative pronoun and complementizer in the above English structures, we can now ask where the functionally equivalent morphology of Austronesian languages fits in, if at all.
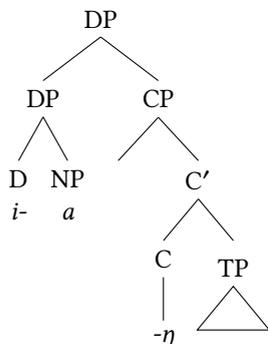
Adelaar (1992) argues convincingly that the *ya* element in *yang* is cognate with the third person singular pronoun *ia* and that the following velar nasal is cognate with the Philippine linker, which we can treat as a type of complementizer.[26] The pronoun *ia* can furthermore be broken down into a person marking determiner element *i* (Ross 2006), plus *a*, a nominal head, as argued for by Reid (2002). The parts of the Malay/Indonesian

---

[25]Ellen Page, "Time to Thrive" Conference, 14 February 2014, Human Rights Campaign video, 0:26, posted and accessed 15 February 2014. Cited from Beatrice Santorini's doubly filled comp example webpage: http://www.ling.upenn.edu/~beatrice/examples/doublyFilledCompExamples.html.

[26]Reid (2002) argues for a similar analysis of Philippine case markers, in which they consist of a nominal head plus a linker. Translating Reid's proposal to the current framework, a case marker like Tagalog *ang* would have an extremely similar structure to Indonesian *yang*. This opens up a possibility whose implications I cannot fully address here, namely, that every Philippine-type DP is akin to a relative clause headed by a dummy nominal. There is some evidence to recommend such a view. Philippine-type DPs can contain a larger range of syntactic material than might naively be expected from an Indo-European perspective. For example, a case marking determiner can have as part of its complement negation and an independent tense domain, as shown in (i).

relativizer thus fit cleanly into the earlier template motivated by English, as shown in (47).

(47)

```
                    DP
                   /  \
                 DP    CP
                /  \     \
               D   NP     C′
               i-   a    /  \
                        C    TP
                        |    △
                       -ŋ
```

Similarly, Kähler (1974) shows that Ngaju Dayak and Old Javanese recruit demonstratives to play the role of relativizer, which can be located in the same DP occupied by *ia* above.

Other languages make use interrogative elements, which we locate on the left branch of CP. In two pioneering investigations of relative clauses in Indonesian languages, Gonda (1943) and Kähler (1974) note the frequency with which **anu* is used as a relativizer, as in (48).[27]

(48) Sundanese

Moal aja deui hajam (a)nu bisa hibar lapas.
NEG EXT anymore chicken RELT can fly fast

'In no case are there anymore chickens which can fly fast.' (Kähler 1974: 264)

In Sundanese, the relativizer is *anu*, a cognate of what Blust & Trussel (2010+) reconstruct as PMP **a-nu* "thing whose name is unknown, avoided, or cannot be remembered: what?"

(i) Tagalog
Ku~kun-in ko bukas ang hindi mo k<in>ain-∅ kahapon.
IMPRF~take-PV 1S.GEN tomorrow NOM NEG 2S.GEN <BEG>eat-PV yesterday

'I will take tomorrow what you didn't eat yesterday.'

Evidence against treating all DPs as full clauses in Philippine-type languages includes the impossibility of dependent form imperatives in DPs (§3 above) as well as the marked nature of topicalization within DPs. The latter argument, however, is weakened by the fact that relative clauses can also plausibly exclude a position for fronted topics. If a relative clause analysis is justified for Philippine-type DPs, then the typological division between languages like Tagalog and Indonesian would have to be characterized not as Philippine-type languages lacking relative clauses but rather lacking bare noun phrase arguments. Historically speaking, bare noun phrase arguments and dedicated relative clause markers clearly appear to be innovations in languages south of the Philippine area.

[27]Kähler (1974) further notes that it appears impossible to reconstruct a dedicated relativizer with any real time depth. I attribute this here to the fact that such elements are not necessary in Philippine-type languages whose event-denoting predicates are already noun-like and can thus serve as direct complements of determiners.
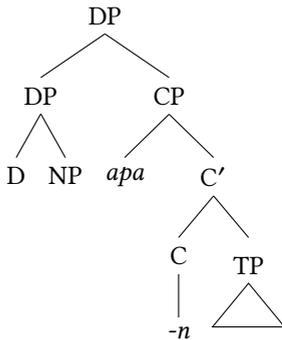
Sangirese *apa(n)*, on the other hand, is cognate with Blust & Trussel's reconstruction of PMP *apa 'what?' and shows evidence for a following nasal linker. The Sangirese relativizer thus fits into our schema as shown in (50).

(49) Sundanese
*I sire **apan** məm-pangasi' su səngkamisa naḷiu e, niuntung bue.*
PM 3P what:LNK AV-plant.rice in one.week past DET lucky EMPH
'They who planted rice last week, are lucky.' (Kähler 1974: 269 citing Adriani 1893-1894)

(50)

```
              DP
           ┌──┴──┐
          DP      CP
         ┌─┴─┐  ┌──┴──┐
        D   NP  apa    C′
                    ┌───┴───┐
                    C        TP
                    │       ╱△╲
                   -n
```

Yet other languages make use of a bleached noun alone. This strategy is extremely common in Sulawesi where we find various derivations of PMP *tau 'person', most often in the reduced form *to*, as in Kulawi (51). The presupposed portion of the clause can be analyzed simply as (52), where all the functional positions are left empty except for the bleached noun.
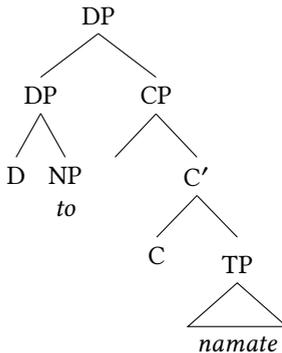
(51) Kulawi (Adriani & Esser 1939: 30)
*Ba bangkele **to** na-mate?*
QM woman RELT PRF-die
'Was it a woman who died?'

(52)

```
              DP
           ┌──┴──┐
          DP      CP
         ┌─┴─┐  ┌──┴──┐
        D   NP         C′
            to      ┌───┴───┐
                    C        TP
                           ╱──△──╲
                           namate
```

Balantak, a language of the Saluan-Banggai subgroup spoken in the eastern side of Central Sulawesi and recently described by van den Berg & Busenitz (2012), displays a fascinating combination of features that put it squarely between Philippine and Indonesian typologies. Like Philippine-type languages, it has a largely intact voice system and the remnants of a case marking system for NPs. The case marker *a* indicates the subject (i.e. the patient of patient voice, agent of agent voice, etc.) when it is post-verbal, as shown in (53). Just as in Tagalog and other Philippine languages, this marker also functions as a definite determiner.

(53) Balantak (van den Berg & Busenitz 2012: 47–48)

    a. *Ma-polos   tuu' **a**   sengke'-ku.*
       INTR.I-hurt very ART back-1S

       'My back really hurts.'

    b. *Boit-i-on   **a**   piso'-muu kabai sobii?*
       sharpen-APP-PV.I ART knife-2P or   let.it.be

       'Should your knife be sharpened or shall we just leave it?'

As in Philippine-type languages, the case marker still allows for complements of all lexical categories, as seen in (54). van den Berg & Busenitz (2012) term such constructions "semi-clefts".
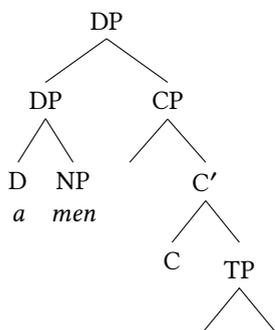
(54) *…raaya'a **a**   mam-bayar.*
     3P   ART AV.I-pay

   '…they were the ones who paid.' (van den Berg & Busenitz 2012: 50)

Remarkably, Balantak has also developed a relative marker *men* from the bleached noun *mian* 'person' (adding further support to the etymology *to* RELT < **tau* 'person' in other languages of Sulawesi). This is seen in (55), where both the case marking determiner *a* and the relativizer *men* co-occur in the presupposition of question.

(55) *Ai   ime **a**   men mae'?*
     EMPH.ART who ART RELT go

   'Who is going?' (van den Berg & Busenitz 2012: 50)

The functional structure of clefts in Balantak, shown in (56), would thus look not very different from Malay/Indonesian.

(56)

What is unique in Balantak is that the determiner in this structure maintains a robust NP case-marking function and can attach directly to verbs in many contexts. Balantak thus offers us a live view of what must have happened throughout Indonesia. The loss of case marking proceeds hand-in-hand with the rise of relativizers. In Balantak questions, it is still the case marker which is obligatory, not the relativizer, as van den Berg & Busenitz (2012) show explicitly in (57).

(57) Balantak

    a. *Pi-pii   takalan a   (men) ala-on-muu?*
        RED-how.many liter   ART RELT take-PV.I-2P

       'How many liters will you take?'

    b. **Pi-pii   takalan (men) ala-on-muu?*
        RED-how.many liter   RELT take-PV.I-2P

       'How many liters will you take?'

But the functional scope of the case marker has also clearly shrunk in comparison to typical Philippine-type languages. Specifically, the nominative determiner *a* only occurs post-verbally in Balantak whereas in Philippine languages we find no such restriction. The loss of this domain would have given rise to the need for a relativizer *men* in positions where *a* was no longer licensed.[28]

5 Referentiality and predication

An idea was put forth earlier that the less referential half of a predication is assigned to the predicate position of a clause while the more referential half is packaged as the

---

[28]Like Balantak, Malagasy also instantiates an intermediate position between canonical Philippine-type languages and Indonesian. It is more complex than the languages considered here in possessing a distinct (i) focus complementizer *no*, (ii) relativizer *izay* and (iii) NP marker *ny*. Keenan's (2008) analysis of Malagasy is close to the one advocated here for Philippine-type languages although Paul (2001); Law (2007); Kalin (2009); Pearson (2009); Potsdam (2006b) show that the syntax of Malagasy cleft constructions is clearly non-equational. Potsdam's (2006b: 220–225) examination of the Malagasy CP in clefts is especially relevant here but space considerations preclude a full comparison.

subject, in line with work on copular clauses in English and other well studied languages. We have also seen in the above how the predicate position in Austronesian languages functions as a kind of de facto focus position by virtue of Austronesian languages tending to package presuppositions as subjects. The mechanics of this turn out to contain some surprises.

First, note that a bare predicate phrase in Tagalog, whether it is headed by an entity-denoting, property-denoting or event-denoting word, must precede the subject, as exemplified in (58) and (59).[29] The basic word order in Tagalog and the vast majority of Philippine-type languages is thus regularly described as predicate-initial on this basis.

(58)   a.   *Guro   ako.*
           teacher 1s.NOM
           'I am a teacher.'

      b.   **Ako   guro.*   Cf. *A teacher is me.
            1s.NOM teacher

(59)   a.   *Matangkad ako.*
           tall   1s.NOM
           'I am tall.'

      b.   **Ako   matangkad.*   Cf. *Tall is me.
            1s.NOM tall

A paradox surfaces, however, when both parts of the predication are referential or definite. In such cases, it appears that the *more* referential portion of the predication must be located in the clause-initial predicate position. In a neutral context, *that* fills the subject position in the English translation of (60). In Tagalog, the demonstrative must be positioned in the clause-initial predicate position. In an English copular clause with a pronominal argument and a definite NP, the pronominal argument will be selected as the subject. In Tagalog, the pronominal argument must always be in predicate position when the other argument is definite, as seen in (61).[30]

(60)   a.   *Iyan   ang problema.*
           that.NOM NOM problem

           'That's the problem.' (Lit. 'The problem is that.')

---

[29]Topicalization is possible to achieve the subject initial orders here but it is marked either by the topic marker *ay* or a very clear intonational break after the subject. In short sentences like (62), the intonational break may be more difficult to hear. Speakers seem to agree however that for the order in (b) to be licit, there must be distinct phonological phrases while this is not true for the (a) sentences. It is in fact possible to make the judgments completely unambiguous through the use of clitics. Specifically, we can compare sentences like the following where the second position pronoun has two forms, a long form, *ikaw*, used in predicate position, and a clitic form *=ka*, used for arguments. When the second person is in a predication with a demonstrative, the clitic form is ungrammatical: *Ikaw iyan* 2s.NOM that.NOM 'That's you' versus **Iyan=ka* 'that.NOM=2SG.NOM'. When the demonstrative is topicalized, the second person pronoun retains its predicate form, *Iyan, ikaw* that.NOM 2SG.NOM 'That, is you'.

[30]As noted in fn.13, Kroeger (1993: 148–149) analyzes such constructions as inversions where the first constituent is the subject and the latter constituent is the predicate. All evidence, however, points to the initial constituents in such sentences behaving as predicates, leading Kroeger (2009) to revise his original analysis.

b. **Ang problema iyan.*
   NOM problem that.NOM

(61) a. *Ako ang guro.*
     1S NOM teacher

     'I am the teacher.'

 b. **Ang guro ako.* Cf. **The teacher is me.*
    NOM teacher 1S

In (62), English and Tagalog again agree in placing the demonstrative in the subject position and the first singular pronoun in predicate position.

(62) a. *Ako iyan.*
     1S.NOM that.NOM

     'That's me.'

 b. ?**Iyan ako.*
    that.NOM 1S.NOM

Based on the above data, we can no longer say that Tagalog merely displays the mirror image of the English subject-predicate order. While both Austronesian languages and English enforce a familiarity condition on subjects (see Mikkelsen 2005: chap.8, for a summary of the English facts), there appears to be an additional role for an extended definiteness or animacy hierarchy in Tagalog and other Philippine languages. The involvement of an animacy hierarchy is clear from the following facts. Just like demonstratives, a third person pronoun must be in predicate position if the other half of the predication is definitely determined, as seen in (63). But when a third person pronoun is in competition with a first person pronoun for predicate position, it is the first person which wins, as shown in (64).

(63) a. *Siya ang problema.*
     3S.NOM NOM problem

     'S/he's the problem.' (Lit. 'The problem is s/he.')

 b. **Ang problema siya.*
    NOM problem 3S.NOM

(64) a. *Ako siya.*
     1S.NOM 3S.NOM

     'S/he's me.'

 b. **Siya ako.*
    3S.NOM 1S.NOM

Although space does not permit a full demonstration of all the possible interactions between NP types, the rules follow a slightly modified version of Aissen's (2003: 437) definiteness hierarchy, shown in (65). When both halves of a predication are referential, the portion higher on the scale in (65) will be selected as predicate.

(65) DEFINITENESS/ANIMATENESS HIERARCHY (Silverstein 1976; Aissen 1999; 2003)
local [1/2] person > third person pronouns > demonstratives/proper name > Definite NP > Indefinite Specific NP > Non-Specific

The only real optionality, as indicated by the lack of ranking above, is found with demonstratives and proper names. When these two types are in a predication relation, either order is acceptable, as seen in (66). This can potentially be linked to the ability of proper names in Tagalog to behave like pronominal clitics (Billings 2005).

(66) a. *Iyan   si   Boboy.*
          that.NOM NOM Boboy

          'That's Boboy.'

      b. *Si   Boboy iyan.*
          NOM Boboy that.NOM

          'That's Boboy.'

In predications where the order is fixed by virtue of the definiteness hierarchy, information structure is flexible. For example, the sentence *ako ang guro* 'I am the teacher' can answer both the question in (5) as well as that in (5). This is unusual in Philippine languages, as the clause-initial predicate position is otherwise reserved for the focus of the sentence rather than the presupposition.[31]

(67) *A: Sino   ang guro?*
          who.NOM NOM teacher

      'Who is the teacher?'

      B: Ako   ang guro.
          1S.NOM NOM teacher

      'I am the teacher.'

---

[31]Aldridge (2013) claims that in predications with two definite DPs (two *ang* phrases), the first is always the focus, exemplified with (i). I am not convinced that a focus reading is necessary or even unmarked on the first *ang* phrase in (i.b). Previous authors have disagreed on the pragmatic status of double *ang* phrase predications in Tagalog. Aldridge argues that predicate fronting in Tagalog (to derive the basic word order) is movement to a focus position. My feeling is rather that the focus interpretation of the predicate is a result of packaging presuppositions as definitely determined subjects. Once the presupposition is subtracted, the left-overs in clause-initial position canonically align with the focus. Examples such as (i) are critical to adjudicating between these two analyses but this must be left to further work.

(i)   a. [*Ang lalaki*] *ang na-kita   ng babae.*
             NOM man   NOM NVOL.PV-see GEN woman

          'The man is who the woman saw.'

      b. [*Ang na-kita   ng babae*] *ang lalaki.*
             NOM NVOL.PV-see GEN woman   NOM man

          'The man is who the woman saw.'

(68) *A: Sino=ka?*
     who.NOM=2S.NOM

    'Who are you?'

    B: Ako   ang guro.
     1S.NOM NOM teacher

    'I am <u>the teacher</u>.'

I would like to offer a potential solution to the paradox of why it is the more definite or referential element that becomes the predicate when both elements are referential, in stark contrast to the canonical packaging of new information as predicate. The pattern can be accounted for by viewing it as the product of two potentially conflicting constraints. On one hand, presuppositions are packaged as *ang* phrases and what is left in the clause-initial position is the de facto focus. The only principle that predicate selection in the strict sense takes into account is whether an element is definite or not. If one element is definite and the other is not the story ends there; the definite element is packaged as subject while the remainder is placed in predicate position. If both elements are definite, another principle comes into play which only relates secondarily to the subject-predicate relation. This principle demands that elements higher on the definiteness/animacy hierarchy *linearly precede* those which are lower on the hierarchy. The clause-initial predicate position is then pressed into service to make the more animate element precede the less animate one.

Several pieces of evidence from other Austronesian languages support this analysis. First of all, as discussed in Kaufman (2014), many Indonesian languages have independently arrived at a split proclitic/enclitic system for agent marking.[32] In all attested examples, third person markers procliticize only if the local persons [1/2] have procliticized. First person furthermore tends to procliticize before second person. This can be seen clearly in the languages of Sumatra, as shown in Table 3 and equally compelling evidence is found in the languages of Sulawesi. On one end of the spectrum, all pronominal agents were enclitic in Old Malay. On the other side of the spectrum, Minangkabau, all such agents are expressed as proclitics. In between, Karo Batak, Gayo and Classical Malay which show a development that respects the animacy hierarchy such that the agents higher on the hierarchy must precede those which are lower.

In an independent development in several languages of Mindanao in the Philippines, the animacy hierarchy also determines the order of clitics within a clitic cluster (Billings & Kaufman 2004; Kaufman 2010a). For instance, in Maranao, a first person clitic always precedes a second person clitic and both first and second person clitics precede third person clitics, as seen in (69).

---

[32]Split proclitic/enclitic patterns in the languages of Sulawesi are argued by van den Berg (1996) to have developed from a full proclitic pattern and by Himmelmann (1996) from a full enclitic paradigm. The history and typology of pronominal proclisis is further discussed by Wolff (1996); Mead (2002); Kikusawa (2003); Billings & Kaufman (2004). I believe the comparative evidence points very clearly to split-paradigms resulting from partial accretion rather than loss, besides the obvious preference of Occam's razor for such an account, but the details do not concern us here.

Table 3: Person marking in the patient voice (Kaufman 2014)

	Old Malay	Karo Batak	Gayo	Clas. Malay	Minangkabau
1SG.	ni-V-**(ŋ)ku**	**ku-V**	**ku-V**	**ku-V**	**den-V**
2SG.	(ni-V-**māmu**)	i-V-**əŋkō**	i-V-**kō**	**kau-V**	**aŋ-V**
3SG.	ni-V-**ña**	i-V-**na**	i-V-**é**	di-V-**ña**	**iño-V**
1PL.EXCL	?	i-V-**kami**	**kami-V**	**kami-V**	**kami-V**
1PL.INCL	ni-V-**(n)ta**	**si-V**	**kitö-V**	**kita-V**	**kito-V**
2PL.	ni-V-**māmu**	i-V-**kam**	i-V-**kam**	**kamu-V**	**kau-V**
3PL.	ni-V-**(n)da**	i-V-**na**	i-V-**é**	di-V-**mereka**	**iño-V**

(69) Maranao (Kaufman 2010a)

     a. *M<iy>a-ilay=ako=ngka.*
         <PRF>PV.NVOL-see=1S.NOM=2S.GEN
         'You saw me.'

     b. *M<iy>a-ilay=ngka=siran.*
         <PRF>PV.NVOL-see=2S.GEN=3P.NOM
         'You saw them.'

Both of these phenomena offer support for the idea that there is an earliness principle at play which makes use of the definiteness/animacy hierarchy. A prediction of this analysis, which is driven by linear precedence, is that no subject-predicate paradox of the type found in Tagalog should exist in Austronesian languages with basic SVO word order. This is because the argument which is higher on the definiteness/animacy hierarchy will both make for a more natural subject and naturally precede the predicate in such languages. This prediction is at least borne out in Indonesian. As seen in (70), even a subject low on the animacy/definiteness hierarchy precedes the predicate in the unmarked word order. In a copular sentence such as that in (71), where a first person pronoun is in a predication relation with a definite NP, the pronouns still takes the canonical subject position. Unlike Tagalog, it cannot felicitously be positioned in predicate position without special topic-focus intonation.

(70) Indonesian
     Serigala bisa membunuh orang.
     wolf   can   AV:kill   person
     'Wolves can kill people'

(71) Indonesian
     a. *Aku guru-nya.*
         1S   teacher-DEF
         'I'm the teacher.'

b. *#Guru-nya   aku.*
   teacher-DEF 1S

Unfortunately, this topic has been left almost completely unexplored for other languages of Indonesia and so it is not yet possible to compare SVO languages of Indonesia with predicate-initial ones more broadly. The predictions of the current analysis are clear though that the unexpected inversions found in Tagalog should only occur in predicate-initial languages.

6 Conclusion

I have explored here several related aspects of predication and information structure in Austronesian languages. I began by arguing for a monoclausal analysis of apparent clefts in Philippine-type languages and tying this to the nominal nature of Philippine-type verbs. I then showed how true biclausal clefts emerge in Indonesian languages where the noun-verb contrast is more robust. In such languages, presupposed verbal material must be relativized before it can occupy subject position. While Indonesian relativizers come from varied sources (bleached nouns, interrogatives, pronouns in combination with the linker), it was shown that all patterns under examination fit neatly into a common syntactic template. Finally, I made an attempt at solving a paradox in the subject-predicate relation of Philippine-type languages. I argued that in addition to a canonical familiarity condition on subjects, there exists a linearity condition which requires that the part of a predication which is higher on the definiteness hierarchy precede the part which is lower. The prediction, which requires further exploration, is that SVO languages should not display these unexpected inversions.

It perhaps deserves emphasizing here that syntacticians have been too hasty in positing English-like constituency structures and lexical categories in the analysis of Austronesian languages. Consequently, important differences between Philippine-type and non-Philippine-type Austronesian languages have been masked. By stepping back from these assumptions, we can begin to explore fundamental problems in the relation between predication and information structure. Although the present work has only scratched the surface, it has hopefully opened a path for further research in how this relation varies across Austronesian languages. The resolution of this problem in Austronesian may very well contribute to answering the philosophical questions around predication first put forth by Plato and Aristotle over two millennia ago and debated today.

Abbreviations

APPL	applicative	BEG	begun aspect
ART	article	COMP	complementizer
AV	actor voice	COP	copula

DET	determiner		NVOL	non-voluntary mood
EMPH	emphatic		OBL	oblique case
EXT	existential		PM	personal marker
GEN	genitive case		PRF	perfective aspect
IMPRF	imperfective aspect		PST	past tense
INTR	intransitive		PV	patient voice
LNK	linker		QM	question marker
NEG	negation		RED	reduplication
NOM	nominative case		RELT	relative marker

References

Abbott, Barbara. 2000. Presuppositions as nonassertions. *Journal of Pragmatics* 32. 1419–1437.

Adelaar, K. Alexander. 1992. *Proto Malayic: The reconstruction of its phonology and parts of its lexicon and morphology*. Canberra: Dept. of Linguistics, Research School of Pacific Studies, the Australian National University.

Adriani, Nicolaus. 1893-1894. Sangireesche teksten. *Bijdragen tot de Taal-, Land- en Volkenkunde van Nederlandsch-Indië* 41/43. 1–168, 383–451, 461–524.

Adriani, Nicolaus & Samuel Jonathan Esser. 1939. *Koelawische taalstudien: Overzicht der spraakkunst, gesprekken en verhalen met vertaling*. Vol. 1. Bandoeng: Nix.

Aissen, Judith. 1999. Markedness and subject choice in optimality theory. *Natural Language and Linguistic Theory* 17. 673–711.

Aissen, Judith. 2003. Differential object marking: Iconicity vs. Economy. *Natural Language and Linguistic Theory* 21. 435–483.

Aldridge, Edith. 2002. Wh-movement in Seediq and Tagalog. In Andrea Rackowski & Norvin Richards (eds.), *Proceedings of AFLA 8: The 8th Meeting of the Austronesian Formal Linguistics Association*, vol. 44, 1–28. Cambridge, MA: MIT Press. MIT Working Papers in Linguistics.

Aldridge, Edith. 2004. *Ergativity and word order in Austronesian languages*. Cornell University (Doctoral dissertation).

Aldridge, Edith. 2013. Wh-clefts and verb-initial word order in Austronesian languages. *Linguistik Aktuell* 208. 71–96.

Ball, Catherine N. 1977. Th-clefts. *Pennsylvania Review of Linguistics* 2. 57–64.

Bauer, Winifred. 1991. Maori *ko* again. *Te Reo* 24. 31–36.

Bauer, Winifred. 1993. *Maori*. London: Routledge.

Besnier, Niko. 2000. *Tuvaluan*. London: Routledge.

Biggs, Bruce. 1974. Some problems of Polynesian grammar. *Journal of the Polynesian Society* 83. 401–426.

Billings, Loren A. 2005. Ordering clitics and postverbal R-expressions in Tagalog: A unified analysis? In Andrew Carnie, Heidi Harley & Sheila Ann Dooley (eds.), *Verb first: On the syntax of verb-initial languages*, vol. 73 (Linguistik Aktuell), 303–339. Amsterdam: John Benjamins Publishing Company.

Billings, Loren A. & Daniel Kaufman. 2004. Towards a typology of Austronesian pronominal clisis. In Paul Law (ed.), *Proceedings of AFLA 11* (ZAS Papers in Linguistics 34), 15–29. Berlin: Zentrum für Allgemeine Sprachwissenschaft.

Bloomfield, Leonard. 1917. *Tagalog texts*. Urbana: University of Illinois Press.

Blust, Robert. 2002. Notes on the history of focus in Austronesian languages. *The History and Typology of Western Austronesian Voice Systems* 518. 63–80.

Blust, Robert. 2015. Case markers of Proto-Austronesian. *Oceanic Linguistics* 54. 436–491.

Blust, Robert & Stephen Trussel. 2010+. *The Austronesian comparative dictionary*. http://www.trussel2.com/ACD.

Byma, Gary F. 1986. *Government and binding in Tagalog: An ergative analysis*. Calgary, Alberta: University of Calgary. (MA thesis).

Capell, Arthur. 1964. Verbal systems in Philippine languages. *Philippine Journal of Science* 93. 231–249.

Chang, Melody Y. 2000. On Tsou wh-questions: Movement or in situ? *Language and Linguistics* 1. 1–18.

Chung, Sandra. 2010. Six arguments for wh-movement in Chamorro. In Donna B. Gerdts, John C. Moore & Maria Polinsky (eds.), *Hypothesis A/Hypothesis B: Linguistic explorations in honor of David M. Perlmutter*, 49th edn. (Current Studies in Linguistics), 91–110. Cambridge, MA: MIT Press.

Cole, Peter & Gabriella Hermon. 2000. Partial wh-movernent: Evidence from Malay. In Uli Lutz, Gereon Müller & Arnim von Stechow (eds.), *Wh-scope marking* (Linguistik Aktuell), 101–130. Amsterdam & Philadelphia: John Benjamins Publishing Company.

Cole, Peter, Gabriella Hermon & Norhaida Aman. to appear. Clefted questions in Malay. *Malay/Indonesian Linguistics*.

Cole, Peter, Gabriella Hermon & Yassir Tjung. 2005. How irregular is wh in situ in Indonesian? *Studies in Language* 29. 553–581.

Constantino, Ernesto. 1965. The sentence patterns of twenty-six Philippine languages. *Lingua* 15. 71–124.

Curme, G. O. 1912. A history of English relative constructions. *The Journal of English and Germanic Philology* 11. 355–380.

Custis, Tonya. 2004. *Word order variation in Tongan: A syntactic analysis*. University of Minnesota. (Doctoral dissertation).

Davidson, Donald. 2005. *Truth and predication*. Cambridge, Massachusetts & London, England: The Belknap Press of Harvard University Press.

den Dikken, Marcel. 2006. *Relators and linkers*. Cambridge, MA: MIT Press.

DeWolf, Charles M. 1988. Voice in Austronesian languages of the Philippine type: Passive, ergative, or neither? In Masayoshi Shibatani (ed.), *Passive and voice*, vol. 16 (Typological Studies in Language), 143–193. Amsterdam & Philadelphia: John Benjamins Publishing Company.

Dukes, Michael. 1998. Evidence for grammatical functions in Tongan. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG98 Conference*. Stanford, CA: CSLI Publications.

Egerod, Soren. 1988. Thoughts on transitivity. *Bulletin Academia Sinica* 59. 369–384.

Foley, William A. 2008. The place of Philippine languages in a typology of voice systems. In Peter K. Austin & Simon Musgrave (eds.), *Voice and grammatical relation in Austronesian languages* (Studies in Constraint Based Lexicalism), 22–44. Stanford: CSLI Publications.

Foley, William A. & Robert D. Van Valin. 1984. *Functional syntax and universal grammar*. Vol. 38 (Cambridge Studies in Linguistics). Cambridge: Cambridge University Press.

Frege, Gottlob. 1892/1997. On concept and object. *The Frege Reader*. 181–193.

Georgopoulos, Carol. 1991. *Syntactic variables*. Dordrecht: Kluwer Academic Publishers.

Gil, David. 1993. Syntactic categories in Tagalog. In Sudapom Luksaneeyanawin (ed.), *Pan-Asiatic Linguistics: Proceedings of the Third International Symposium on Language and Linguistics, Bangkok, January 8–10, 1991*, 1136–1150. Bangkok: Chulalongkorn University Press.

Gonda, J. 1943. Indonesischerelativa. *Bijdragen tot de Taal-, Land- en Volkenkunde van Nederlandsche-Indig* 102. 501–537.

Heycock, Caroline. 2013. The syntax of predication. In Marcel Den Dikken (ed.), *The Cambridge handbook of generative syntax*, 322–352. Cambridge: Cambridge University Press.

Himmelmann, Nikolaus P. 1986. *Morphosyntactic predication: A functional-operational approach*. Vol. 62 (AKUP - Arbeiten des Kölner Univeralien Projekts). Cologne: Institut für Sprachwissenschaft der Universität zu Köln.

Himmelmann, Nikolaus P. 1987. *Morphosyntax und Morphologie. Die Ausrichtungsaffixe im Tagalog*. Vol. 8 (Studien zur theoretischen Linguistik). München: Wilhelm Fink.

Himmelmann, Nikolaus P. 1991. *The Philippine challenge to universal grammar*. Vol. 15. Cologne: Institut für Sprachwissenschaft der Universität zu Köln.

Himmelmann, Nikolaus P. 1996. Person marking and grammatical relations in Sulawesi. *Papers in Austronesian Linguistics* 84. 115–136. No. 3, Series A.

Himmelmann, Nikolaus P. 1997. *Deiktikon, Artikel, Nominalphrase: Zur Emergenz syntaktischer Struktur*. Vol. 362 (Linguistische Arbeiten). Tübingen: Niemeyer.

Himmelmann, Nikolaus P. 2005. Typological characteristics. In K. Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar*, 110–181. London: Routledge.

Jespersen, Otto. 1965. *The philosophy of grammar*. New York: W. W. Norton & Company.

Jespersen, Otto. 1937/1984. *Analytic syntax*. Chicago: University of Chicago Press.

Kadmon, Nirit. 2000. *Formal pragmatics*. Oxford: Blackwell Publishers.

Kähler, Hans. 1974. Relative clause formation in some Austronesian languages. *Oceanic Linguistics* 13. 257–277.

Kalin, Laura. 2009. News about the *no*-phrase: Specificational pseudoclefts in Malagasy. In Sandy Chung, Daniel Finer, Ileana Paul & Eric Potsdam (eds.), *Proceedings of the sixteenth meeting of the Austronesian Formal Linguistics Association (AFLA)*, 61–75.

Kaufman, Daniel. Forthcoming. Response to commentators. *Theoretical Linguistics*.

Kaufman, Daniel. 2005. Aspects of pragmatic focus in Tagalog. In I Wayan Arka & Malcolm Ross (eds.), *The many faces of Austronesian voice systems: Some new empirical*

studies (number 571 in Pacific Linguistics), 175–196. Canberra: Research School of Pacific & Asian Studies, ANU.

Kaufman, Daniel. 2009a. Austronesian nominalism and its consequences: A Tagalog case study. *Theoretical Linguistics* 35. 1–49.

Kaufman, Daniel. 2009b. Austronesian typology and the nominalist hypothesis. In K. Alexander Adelaar & Andrew Pawley (eds.), *Austronesian historical linguistics and culture history: A festschrift for Robert Blust* (Pacific Linguistics), 197–226. Canberra: ANU Press.

Kaufman, Daniel. 2010a. The grammar of clitics in Maranao. In Loren A. Billings & Nellke Goudswaard (eds.), *Piakandatu ami dr. Howard p. Mckaughan*, 132–157. Manila: Linguistic Society of the Philippines & SIL.

Kaufman, Daniel. 2010b. *The morphosyntax of Tagalog clitics: A typological approach.* Ithaca, NY: Cornell University (Doctoral dissertation).

Kaufman, Daniel. 2014. The syntax of Indonesian imposters. In Chris Collins (ed.), *Cross-linguistic studies of imposters and pronominal agreement* (Oxford Studies in Comparative Syntax), 89–120. New York: Oxford University Press.

Kaufman, Daniel. 2017. Lexical category and alignment in Austronesian. In Lisa Travis, Jessica Coon & Diane Massam (eds.), *Oxford handbook of ergativity*, 589–630. Oxford: Oxford University Press.

Keenan, Edward L. 1976. Towards a universal definition of "subject". In Charles Li & Sandra A. Thompson (eds.), *Subject and topic*, 303–333. New York: Academic.

Keenan, Edward L. 2008. Voice and relativization without movement in Malagasy. *Natural Language and Linguistic Theory* 26. 467–497.

Kikusawa, Ritsuko. 2003. The development of some Indonesian pronominal systems. In *Historical Linguistics 2001: Selected Papers from the 15th International Conference on Historical Linguistics, Melbourne, 13–17 August 2001*, 237–268. John Benjamins Publishing Company.

Kroeger, Paul. 1993. *Phrase structure and grammatical relations in Tagalog.* Stanford: Center for the Study of Language & Information, Stanford University.

Kroeger, Paul. 2007. McKaughan's analysis of Philippine voice. In Loren A. Billings & Nelleke Goudswaard (eds.), *Piakandatu ami Dr. Howard P. McKaughan*, 41–46. Manila: Linguistic Society of the Philippines & SIL Philippines.

Kroeger, Paul. 2009. Malagasy clefts from a Western Malayo-Polynesian perspective: Commentary on the paper by Hans-Martin Gärtner. *Natural Language and Linguistic Theory* 27. 817–838.

Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus and the mental representations of discourse referents.* Cambridge, UK: Cambridge University Press.

Lambrecht, Knud. 2001. A framework for the analysis of cleft constructions. *Linguistics* 39. 463–516.

Law, Paul. 2007. The syntactic structure of the cleft construction in Malagasy. *Natural Language and Linguistic Theory* 25. 765–823.

Lemaréchal, Alain. 1991. Dérivation et orientation dans les langues des Philippines. *Bulletin de la Société de Linguistique de Paris* 86. 317–358.

Lopez, Cecilio. 1937/1977. Preliminary study of the affixes in Tagalog. *Selected writings of Cecilio Lopez in Philippine linguistics* 1. 28–104.

Mead, David. 2002. Proto-Celebic focus revisited. In Malcolm Wouk & Fay Ross (eds.), *The history and typology of Western Austronesian voice systems* (Pacific Linguistics), 143–177. Canberra: Research School of Pacific Asian Studies, ANU.

Mikkelsen, Line. 2005. *Copular Clauses: Specification, predication and equation.* Vol. 85 (Linguistik Aktuell). Amsterdam & Philadelphia: John Benjamins Publishing Company.

Mikkelsen, Line. 2011. Copular clauses. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics: An international handbook of natural language meaning*, vol. 2, 1805–1829. Berlin: Mouton de Gruyter.

Modrak, D. K. 1985. Forms and compounds. In James Bogen & James E. McGuire (eds.), *How things are: Studies in predication and the history of philosophy and science*, 85–99. Dordrecht, Boston & London: D. Reidel Publishing Company.

Nagaya, Naonori. 2007. Information structure and constituent order in Tagalog. *Language and Linguistics* 8. 343–372.

Naylor, Paz B. 1975. Topic, focus, and emphasis in the Tagalog verbal clause. *Oceanic Linguistics* 14. 12–79.

Naylor, Paz B. 1980. Linking, relation-marking, and Tagalog syntax. In Paz B. Naylor (ed.), *Austronesian Studies, papers from the 2nd Eastern Conference on Austronesian Languages*, 33–49. Ann Arbor: The University of Michigan.

Naylor, Paz B. 1995. Subject, topic, and Tagalog syntax. In D. Benett, T. Bynon & G. B. Hewitt (eds.), *Subject, voice and ergativity*, 161–201. London: SOAS.

Otsuka, Yuko. 2000. *Ergativity in Tongan.* University of Oxford (Doctoral dissertation).

Paul, Ileana. 2000. *Malagasy clause structure.* McGill University (Doctoral dissertation).

Paul, Ileana. 2001. Concealed pseudo-clefts. *Lingua* 111. 707–727.

Paul, Ileana. 2008. On the topic of pseudoclefts. *Syntax* 11. 91–124.

Pearson, Matthew. 2009. Another look at *no*: Pseudo-clefts and temporal clauses in Malagasy. In Sandy Chung, Daniel Finer, Ileana Paul & Eric Potsdam (eds.), *Proceedings of the sixteenth meeting of the Austronesian Formal Linguistics Association (AFLA)*, 165–179.

Potsdam, Eric. 2006a. More concealed pseudoclefts in Malagasy and the clausal typing hypothesis. *Lingua* 116. 2154–2182.

Potsdam, Eric. 2006b. The cleft structure of Malagasy wh-questions. In Hans-Martin Gärtner, Paul Law & Joachim Sabel (eds.), *In clause structure and adjuncts in Austronesian languages*, 195–232. Berlin: Mouton de Gruyter.

Potsdam, Eric. 2009. Austronesian verb-initial languages and *wh*-question strategies. *Natural Language and Linguistic Theory* 27. 737–771.

Rackowski, Andrea. 2002. *The structure of Tagalog: Specificity, voice and the distribution of arguments.* Cambridge, MA: MIT (Doctoral dissertation).

Reeve, Matthew. 2010. *Clefts.* University College London (Doctoral dissertation).

Reid, Lawrence Andrew. 2002. Determiners, nouns, or what? Problems in the analysis of some commonly occurring forms in Philippine languages. *Oceanic Linguistics* 41. 295–309.

Richards, Norvin. 1998. Syntax versus semantics in Tagalog wh-extraction. *UCLA occasional papers in linguistics* 21. 259–275.

Richards, Norvin. 2000. Another look at Tagalog subjects. In Ileana Paul, Vivianne Phillips & Lisa Travis (eds.), *Formal issues in Austronesian linguistics*, 105–116. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Richards, Norvin. 2009a. Nouns, verbs, and hidden structure in Tagalog. *Theoretical Linguistics* 35. 139–152.

Richards, Norvin. 2009b. The Tagalog copula. In Sandy Chung, Daniel Finer, Ileana Paul & Eric Potsdam (eds.), *Proceedings of the sixteenth meeting of the Austronesian Formal Linguistics Association (AFLA)*, 181–195. London: AFLA Proceedings Project, University of Western Ontario.

Ross, Malcolm. 2002. The history and transitivity of western Austronesian voice and voice marking. In Fay Wouk & Malcolm Ross (eds.), *The history and typology of Western Austronesian voice systems* (Pacific Linguistics), 17–63. Canberra: Research School of Pacific & Asian Studies, ANU.

Ross, Malcolm. 2006. Reconstructing the case-marking and personal pronoun systems of Proto-Austronesian. In Yongli Zhang (ed.), *Streams converging into an ocean: Festschrift in honor of Professor Paul Jen-kuei Li on his 70th birthday*, vol. 5 (Language and Linguistics Monographs), 521–563. Taipei: Institute of Linguistics, Academica Sinica.

Ross, Malcolm. 2009. Proto-Austronesian verbal morphology: A reappraisal. In K. Alexander Adelaar & Andrew Pawley (eds.), *Austronesian historical linguistics and culture history: A festschrift for Robert Blust* (Pacific Linguistics), 295–326. Canberra: Research School of Pacific & Asian Studies, ANU.

Ross, Malcolm. 2015. Reconstructing Proto-Austronesian verb classes. *Language and Linguistics* 13. 279–315.

Schachter, Paul. 1976. The subject in Philippine languages: Topic, actor, actor- topic, or none of the above. In Charles Li (ed.), *Subject and topic*, 491–518. New York: Academic Press.

Schachter, Paul & Fe T. Otanes. 1982. *Tagalog reference grammar*. Berkeley, CA: University of California Press.

Scheerer, Otto. 1924. On the essential difference between the verbs of the European and the Philippine languages. *Philippine Journal of Education* 7. 1–10.

Seiter, William J. 1975. *Information questions in Philippine languages*. San Diego: University of California.

Seiter, William J. 1980. *Studies in Niuean syntax*. New York: Garland Publishingo.

Shibatani, Masayoshi. 2009. Elements of complex structures, where recursion isn't: The case of relativization. In Talmy Givon & Masayoshi Shibatani (eds.), *Syntactic complexity: Diachrony, acquisition, neuro-cognition, evolution*, vol. 85 (Typological Studies in Language), 163–198. Amsterdam: John Benjamins Publishing Company.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In R. M. W. Dixon (ed.), *Grammatical categories in Australian languages*, 112–171. Canberra: Australian Institute of Aboriginal Studies.

Starosta, Stanley, Andrew K. Pawley & Lawrence Reid. 1982. The evolution of focus in Austronesian. In Stephen A. Wurm & Lois Carrington (eds.), *Papers from the Third International Conference on Austronesian Linguistics, Vol.2. Tracking the travellers*, 145–170. Canberra: Pacific Linguistics.

Tchekhoff, Claude. 1981. *Simple sentences in Tongan.* Vol. 81. Canberra: ANU Press.

van den Berg, René. 1996. The demise of focus and the spread of conjugated verbs in sulawesi. *Papers in Western Austronesian Linguistics* 3. 89–114.

van den Berg, René & Robert I. Busenitz. 2012. *A grammar of Balantak: A language of Eastern Sulawesi.* Hein Steinhauer (ed.). Vol. 40 (SIL eBook). SIL International.

Williams, Edwin. 1997. The asymmetry of predication. *Texas Linguistic Forum* 38. 323–333.

Wolff, John U. 1973. Verbal inflection in Proto-Austronesian. *Philippine Journal of Linguistics Special Monograph.* Parangal kay Cecilio Lopez 4. 71–91.

Wolff, John U. 1996. The development of the passive verb with pronominal prefix in Western Austronesian langauges. In Bernd Nothofer (ed.), *Reconstruction, Classification, description: Festschrift in honor of Isidore Dyen*, 15–40. Hamburg: Abera.

Yap, Foong Ha. 2011. Referential and non-referential uses of nominalization construction in Malay. In Foong Ha Yap & Janick Wrona (eds.), *Nominalization in Asian languages: Diachronic and typological perspectives* (Typological Studies in Language), 627–658. Amsterdam/Philadelphia: John Benjamins Publishing Company.