Bikol clefts and topics and the Austronesian extraction restriction

Michael Yoshitaka Erlewine and Cheryl Lim, National University of Singapore

May 2022

Many Austronesian languages exhibit an extraction restriction whereby only one particular DP — the "pivot" argument, the choice of which is reflected by morphology on the verb — can be Ā-extracted. We show that such extraction restrictions can vary between different Ā-constructions in Bikol: local clefting is limited to the pivot, whereas topicalization can target pivots and non-pivot agents but not non-pivot themes of transitive verbs. Following the phase-theoretic, locality-based approach to such extraction asymmetries in related Austronesian languages, we propose that clefting and topicalization differ in the featural specifications of their probes, with clefting necessarily targeting the closest DP and topicalization simply seeking the closest topic constituent. Evidence for this approach comes from the behavior of long-distance clefting, which may target certain non-pivot arguments and involve gaps or resumptive pronouns. The inventory of different long-distance cleft types is explained by the possibility of embedded topicalization and hanging topic left dislocation feeding higher clefting. Our study strengthens the view that the classic Austronesian pivot-only extraction restriction is best characterized in terms of syntactic locality, rather than as a restriction on the grammatical function or morphological case of movement targets.

Contents

1 Introduction

Many Austronesian languages of the Philippines and beyond exhibit a so-called "voice system" or "Philippine alignment"; in such languages, each clause has one argument that we call the "pivot" and, among other properties, this designated pivot argument is the only DP that can be targeted for Ā-extraction. The nature of this extraction restriction — which has also been described by some authors as a "subject-only" or "absolutive-only" restriction — has been a focal point for typological and theoretical discussions of extraction asymmetries (Keenan and Comrie, 1977; Schachter and Otanes, 1972; Aldridge, 2004; Rackowski and Richards, 2005, a.o.) and is also central to discussions of the notion of "subjecthood" in Austronesian and beyond (Keenan, 1976; Schachter, 1976, 1996; Shibatani, 1988; Guilfoyle, Hung, and Travis, 1992; Kroeger, 1993, a.o.).

In this paper, we describe patterns of Ā-extractions — specifically, clefting and topicalization — in Bikol, an Austronesian language of the central Philippines closely related to Tagalog.[1] At first glance, Bikol exhibits a familiar Philippine voice system. In example (1a), the theme *lalaki* 'man' has been chosen as the designated pivot and therefore is in nominative case. Patient Voice morphology on the verb reflects that the nominative argument is the verb's theme. Local clefting is limited to this pivot argument, as in (1b,c), and must leave a post-verbal gap, as we show below. Clefting of the non-pivot agent *eskwela* 'student' in (1c) is ungrammatical both when retaining its original genitive case marker or changing it to nominative case. Local clefting thus manifests the basic pivot-only extraction asymmetry predicted of Philippine voice system languages.

(1)   **Local clefting is limited to the pivot:**

    a.   <u>Baseline Patient Voice (PV) clause:</u>

        G\<in\>adan su   lalaki kaso eskwela.

        PV-kill   NOM man GEN student

        'The student killed the man.'

    b.   <u>Grammatical pivot (theme) cleft:</u>

        Su   lalaki su   [g\<in\>adan ____ kaso eskwela].

        NOM man NOM PV-kill   GEN student

        'It's the man that the student killed.'

---

[1] Glosses follow the Leipzig conventions, with the following additions: AV = Actor Voice, PV = Patient Voice, LV = Locative Voice, BV = Benefactive Voice; see §2. All uncredited data come from our elicitation with two speakers of Bikol in Singapore, led by the second author.

c.   <u>Ungrammatical non-pivot agent cleft:</u>

        *Su   / kaso eskwela su   [g\<in\>adan su   lalaki ___].

        NOM / GEN student NOM PV-kill   NOM man

        Intended: 'It's the student that killed the man.'

In contrast, we observe that local topicalization in Bikol can target both pivots and non-pivot agents. Examples (2a) and (2b) are both grammatical and express the same proposition that 'The student killed the man.' In (2a), the pivot *lalaki* 'man' is topicalized to a pre-verbal position, whereas in (2b), the non-pivot agent *eskwela* 'student' is topicalized to a pre-verbal position. When the non-pivot agent is topicalized in (2b), its case marking changes to be in nominative case, resulting in a clause with two nominative phrases; however, (2b) unambiguously means 'The student killed the man,' and not 'The man killed the student.'

(2)   **Grammatical topicalization of theme pivot and non-pivot agent:**

    a.   Su   lalaki [g\<in\>adan ___ kaso eskwela].

        NOM man   PV-kill   GEN student

    b.   Su   eskwela [g\<in\>adan su   lalaki ___].

        NOM student   PV-kill   NOM man

        'The student killed the man.'

The availability of non-pivot agent topicalization in (2b) is surprising against the backdrop of the widely-discussed pivot-only restriction on Ā-extraction in these languages.[2] In addition, the availability of two nominative-marked arguments in (2b) raises questions for the nature of nominative case and the interaction of voice marking and case in these languages, which we will address.

    The core of our proposal will be that clefting and topicalization involve probes with different featural specifications: Clefting involves a head that probes for a focus that is the closest DP (see Branan and Erlewine, to appearb) — echoing Aldridge's (2004; 2017) proposals of a φ-probe for all Ā-extractions in related voice system languages — whereas topicalization probes straightforwardly for a topic feature. Following the work of Aldridge (2004, 2008), Rackowski and Richards (2005), and others, the pivot argument in Austronesian voice system languages is the highest argument in the lower phase, in a (outer) specifier of *v*P. Due to their differing specifications, clefting cannot attract another DP past the pivot, whereas topicalization can skip the highest DP (the pivot) and attract a non-pivot agent occupying the inner specifier of *v*P. When the agent is

---

[2] Examples similar to (2b) in other, related Philippine languages have been observed in passing by some authors. However, such examples are few and far between and they have not received serious attention. We return to this point in the conclusion, section 6.

itself the pivot, in Actor Voice (AV), it is the only DP at the edge of the *v*P phase. Probing obeys Phase Impenetrability (Chomsky, 2000), explaining the unavailability of non-pivot theme topicalization, in (3).

(3)   **Non-pivot themes cannot be topicalized:**

   *Su   eskwela [nag-gadan su   lalaki ___].
   NOM student   AV-kill   NOM man

   Intended: 'The man killed the student.'

Support for our locality-based approach will come from the behavior of long-distance clefting. In contrast to local clefting which is restricted to pivots, as in (1) above, long-distance clefting can target non-pivot agents, as seen in (4). We argue that such examples involve a step of non-pivot agent topicalization within the embedded CP that makes the embedded non-pivot agent the highest DP within the embedded clause, which then counts as the closest DP target for matrix clefting.

(4)   **Long-distance cleft of embedded non-pivot agent:**

   Su   eskwela su   [pig-balita ning radyo [$_{CP}$ na   g<in>adan su   lalaki ___]].
   NOM student NOM   PV-report GEN radio   that PV-kill   NOM man

   'It's the student that the radio reported *t* killed the man.'

This and additional data inform our description of the nature of the basic Austronesian pivot-only extraction restriction, obeyed in Bikol by local clefting. We argue that the observed "pivot-only" extraction restriction must be characterized in terms of syntactic locality, reflecting the attraction of the closest DP, rather than any requirement to attract pivots or even nominative DPs.

We additionally discuss hanging topic left dislocation (HTLD), a non-movement-derived form of topic with an obligatory corresponding pronoun. As a non-movement construction, HTLD can be used for any DP argument: pivots, non-pivot agents, as well as non-pivot themes of transitive verbs. Just as long-distance clefts can be fed by movement topicalization as in (4), allowing for non-pivot agent clefts, long-distance clefts can be fed by embedded HTLD. This results in long-distance clefts with resumptive pronouns, which have no restriction on the arguments they can target unlike gapped clefts.

This paper is structured as follows. Section 2 discusses the interaction of case marking and voice morphology in Philippine-type languages and how these properties manifest in Bikol. Local Ā-extraction facts are presented in section 3, followed by our core analysis in section 4. In section 5, we discuss long-distance clefting, which will motivate a locality-based characterization of the Austronesian extraction

restriction. Along the way, we describe the two different types of pre-verbal topics, the organization of the *v*P phase edge, and the determination of morphological case in Bikol.

2   Case and voice in Bikol

In this section we introduce basic properties of Bikol morphosyntax that will be relevant for our study. Many Austronesian languages, including Bikol, exhibit a particular constellation of case marking, verbal morphology, and extraction interactions that have been termed a "voice system." A summary of these key properties, building on prior work such as Schachter 1976, 1996, Guilfoyle et al. 1992, and Ross 2009, is quoted here in (5) from Erlewine, Levin, and Van Urk 2017 with minor modification. As is noted in these works, there is significant variation in the terms used for such systems in previous literature.[3]

(5)   **Characteristics of Austronesian-type voice systems:**

(modified from Erlewine et al., 2017: 376)

   a.   A privileged argument: One argument is designated the "pivot," and is realized in a particular morphological form and/or structural position, regardless of its original thematic role or grammatical function.

   b.   Articulated voice morphology: Morphology on the verb varies with the choice of pivot, including options for taking certain oblique arguments as pivot.

   c.   Extraction restriction: Ā-extraction (*wh*-movement, relativization, topicalization, etc.) of DPs is limited to the pivot argument.

   d.   Marking of non-pivot agents: Non-pivot agents are morphologically marked, often coinciding with the form of possessors (i.e. genitive case).

The core voice system properties in (5) are all readily observed in Bikol, although in the rest of the paper we will show that the facts surrounding Ā-extraction (5c) are in reality more complicated. In the rest of this section, properties (5a), (5b), and (5d) of the Bikol voice system will be presented. Data on Ā-extraction which partially supports the characterization in (5c) will be presented in the following section.

---

[3] Other terms for what we call the "pivot" here include terms such as "subject," "trigger," "topic," and "focus." We especially avoid the latter terms, as they suggest particular information-structural notions. As we will see, the pivot argument is often an information-structural topic or focus, but under certain circumstances, non-pivot arguments can be associated with these categories. Many authors refer to the pivot as "subject," while acknowledging that some diagnostics appear to track thematic base positions in many of these languages; see e.g. Shibatani 1988; Kroeger 1991, 1993.

Canonical word order in Bikol is predicate-initial. Consider the examples in (6) below, which all express the basic proposition that 'The woman bought cheese at a store for Andrew.' In each example, there is one pivot DP in nominative case, in bold in (6), and voice morphology on the verb that correlates with this choice of pivot argument. The pivot can be the thematic agent (6a) or theme (6b), but can also be a non-core thematic argument such as a location (6c) or a beneficiary (6d) which is otherwise expressed as an oblique. Post-verbal word order is free; only one word order is given for each example here.

(6)   **Voice alternation in Bikol:**

    a.   Actor Voice (AV):

       Nag-bakal **su**   **babayi** ning keso   sa   tindahan para ki   Andrew.

       AV-buy   NOM woman GEN cheese DAT store   for   DAT Andrew

       'The woman bought cheese at a/the store for Andrew.'

    b.   Patient Voice (PV):

       Pig-bakal kaso babayi **su**   **keso**   sa   tindahan para ki   Andrew.

       PV-buy   GEN woman NOM cheese DAT store   for   DAT Andrew

       'The woman bought the cheese at a/the store for Andrew.'

    c.   Locative Voice (LV):

       Pig-bakal-**an** kaso babayi ning keso   **su**   **tindahan** para ki   Andrew.

       LV-buy   GEN woman GEN cheese NOM store   for   DAT Andrew

       'The woman bought cheese at the store for Andrew.'

    d.   Benefactive Voice (BV):

       I-b<**in**>akal kaso babayi ning keso   sa   tindahan **si**   **Andrew**.

       BV-buy   GEN woman GEN cheese DAT store   NOM Andrew

       'The woman bought cheese at a/the store for Andrew.'

Bikol distinguishes three different cases — nominative, genitive, and dative — with a rich inventory of surface forms that vary based on animacy and number. The table in (7) covers all case marker forms in examples that we will discuss, involving singular noun phrases, as well as the corresponding third-singular animate pronouns and demonstrative pronouns, which are used for inanimate referents. See Mintz 1973: ch. 2 and McFarland 1974 for more detailed descriptions of these inventories. In examples throughout, we will simply gloss these markers as NOM, GEN, or DAT respectively. Genitive and nominative animate pronouns are second position clitics; see also Erlewine and Levin 2021.

(7) **Case markers for singular noun phrases and corresponding pronouns:**

	case markers		pronouns	
	proper name	common noun	3sg animate	demonstrative
nominative	*si*	*su*	*=siya*	*ito*
genitive	*ni*	*kaso/ning*[4]	*=niya*	*kaito*[5]
dative	*ki*	*sa*	*sainya*	*kaito*

Non-pivot core arguments are generally in genitive case. In addition, specific non-pivot themes may appear in dative case rather than genitive case as in (8), but all non-pivot agents are in genitive case.[6]

(8) **Specific non-pivot themes are dative-marked:**

Nag-hiling **sa** babayi si Andrew.

av-see DAT woman NOM Andrew

'Andrew saw the woman.'

Although the voice system allows for different arguments to be the pivot and hence nominative, in the canonical, predicate-initial word order, it is not possible for two arguments of the clause to simultaneously be nominative. This explains the ungrammaticality of (9) below, in contrast to (6b) above. The ungrammaticality of (9) with *sa tindahan* shows that the ungrammaticality of (9) is not due to a surface ban on adjacent nominative phrases.

(9) **Only one (post-verbal) argument may be nominative:**

*Pig-bakal **su** babayi (sa tindahan) **su** keso.

pv-buy NOM woman DAT store NOM cheese

Intended: 'The woman bought the cheese (at the store).'

It's worth noting that this "voice system" descriptively differs from familiar "voice" alternations in European (and other) language families. First, neither the Actor Voice nor Patient Voice appears to be

---

[4] Genitive inanimate common nouns are introduced by *ning*. Genitive animate common nouns can be preceded by *kaso* or *ning*.

[5] *Kaito* is both the genitive and dative form of the demonstrative (Mintz, 2019: 184).

[6] The Tagalog equivalent of this Differential Object Marking is well-studied; see e.g. Schachter and Otanes 1972 and, more recently, Latrouite 2011a, Sabbagh 2016, and Collins 2019. From a broader, pan-Austronesian perspective, Chen 2017 describes such markers as variants of the case marker for non-pivot themes (in our terms, genitive) that encode definiteness or specificity, rather than as a distinct case.

morphologically or syntactically simpler on the surface, leading some authors to refer to such systems as "symmetric" voice systems; see especially Foley 2008. Second, in the Non-Actor Voices (NAV) — which some earlier works describe as "passives" — the agent argument continues to be a DP core argument of the clause, rather than a demoted oblique. The present paper will offer further support for the view that NAV agents are full-fledged DP arguments.

Finally, we note that there is a not insignificant tradition of describing Philippine languages as exhibiting ergative/absolutive alignment. See for example Payne 1982, De Guzman 1988, Gerdts 1988, Mithun 1994, and Aldridge 2004. Under this view, Actor Voice clauses are formally intransitive, with an oblique theme, and Non-Actor Voice clauses are formally transitive. Pivots are absolutive and NAV agents are ergative, with the case on non-specific AV themes then being a homophonous oblique. On this point, see especially Aldridge's (2004; 2012) ergative analysis for Tagalog, whose voice morphology and case facts parallel the Bikol facts above. The pivot-only Ā-extraction restriction is then an absolutive-only extraction restriction, which is also attested in other language families where the "ergative" designation is less controversial, such as Inuit, Mayan, and Salishan; see Deal 2016 and Polinsky 2017 for two recent overviews. This ergative hypothesis for Philippine-type voice system languages has however been controversial; see especially Chen 2017, Erlewine, Levin, and Van Urk 2017, and Kaufman 2017 for recent critical discussion.

In this paper we use the terms nominative and genitive for the two core cases in Bikol, as in the earlier examples in this section, and later present an analysis for Bikol case and voice in these terms. However, the empirical contribution of our paper as well as its theoretical import is logically separable from this choice. Our core proposal for Bikol extraction facts, in section 4, in fact largely follows the syntax for Austronesian voice systems proposed in Aldridge's work. Lessons for the analysis of syntactic ergativity — to the extent that Philippine-type voice system languages should be described as ergative — will be presented at the end of section 5.

3   Local clefts and topics

In this paper we discuss the clefting of DPs and two types of DP topic constructions in Bikol, which we refer to as topicalization and hanging topic left dislocation (HTLD). We limit our attention to dependencies with DPs, as the movement of non-DPs behaves quite differently in Philippine languages.[7] In this section, we specifically consider local clefts and topics. In the interest of space, we will concentrate on extractions

---

[7] See for example the discussion of Tagalog "adjunct fronting" in Kroeger 1991, as well as Hsieh 2020. See also footnote 15 below and brief discussion of pre-verbal PPs in section 4.3.

of agent and theme arguments of notionally transitive verbs from Actor Voice (AV), Patient Voice (PV), and Locative Voice (LV) clauses.

In our work we have also investigated DP *wh*-questions. As in many other Austronesian languages, ex-situ DP *wh*-questions are formally clefts (see Potsdam 2009 for an overview) and indeed the patterns we report here for local clefts (and long-distance clefts below) replicate with DP *wh*-phrases.

3.1 Clefts

As noted above, it is often claimed that only the pivot can be Ā-extracted in voice system languages — famously described as a "subject-only" restriction by Keenan and Comrie (1977) — including in closely related Philippine languages (Kroeger, 1991; Reid and Liao, 2004; Aldridge, 2004; Rackowski and Richards, 2005). This characterization indeed holds for local clefting, our first Ā-construction in Bikol. Clefts have two parts: the exhaustive focus or focus-containing phrase and the background (a gapped clause), separated by a nominative case marker.[8] Example (10) shows that only the agent pivot can be clefted out of an AV clause. Clefting the non-pivot theme in (10b) is ungrammatical, whether retaining the original genitive case marker *ning* or switching to nominative case. Example (11) similarly shows that only the theme pivot can be clefted from a PV clause, as we also saw in (1) above.

(10) **Local clefting from an AV clause:**

    a. <u>Grammatical agent pivot cleft:</u>

        [focus Su   babayi] su   [background nag-kaon (*=siya)   ning/sa   keso   ].

            NOM woman NOM   AV-eat   NOM.3sg GEN/DAT cheese

        'It's the woman that ate (the) cheese.'

    b. <u>Ungrammatical theme cleft:</u>

       *[focus Su/ning/sa   keso   ] su   [background nag-kaon su   babayi].

          NOM/GEN/DAT cheese NOM   AV-eat   NOM woman

        Intended: 'It's (the) cheese that the woman ate.'

---

[8] The fronted constituent in a cleft is also sometimes called a "pivot," but we avoid this term here. "Pivot" in this paper uniformly refers to the argument cross-referenced by voice.

(11)   **Local clefting from a PV clause:**

    a.   <u>Grammatical theme pivot cleft:</u>

        [focus Su   keso] su   [background k<in>aon kaso babayi (*ito)   ].

            NOM cheese NOM   PV-eat   GEN woman   NOM.DEM

        'It's the cheese that the woman ate.'

    b.   <u>Ungrammatical agent cleft:</u>

      *[focus Su/kaso   babayi] su   [background k<in>aon su   keso].

            NOM/GEN woman NOM   PV-eat   NOM cheese

        Intended: 'It's the woman that ate the cheese.'

The ungrammaticality of the nominative pronouns in both examples shows that local clefts must be gapped; i.e. they must have a post-verbal pivot gap, not a corresponding pronoun. Post-verbal gaps will generally not be indicated in examples, due to the flexible post-verbal word order mentioned in section 2.

    We also present clefts from Locative Voice (LV) in (12) below. (12) shows that only the locative pivot can be clefted. Clefting the non-pivot agent as in (12b) or the non-pivot theme as in (12c) is ungrammatical.

(12)   **Local clefting from a LV clause:**

    a.   <u>Grammatical locative pivot cleft:</u>

        [focus Su   tindahan] su   [background pig-bakal-an kaso babayi ning/sa   keso   ].

            NOM store   NOM   LV-buy   GEN woman GEN/DAT cheese

        'It's at the store that the woman bought (the) cheese.'

    b.   <u>Ungrammatical agent cleft:</u>

      *[focus Su/kaso babayi] su   [background pig-bakal-an su   tindahan ning keso   ].

            NOM/GEN woman NOM   LV-buy   NOM store   GEN cheese

        Intended: 'It's the woman that bought cheese at the store.'

    c.   <u>Ungrammatical theme cleft:</u>

      *[focus Su/ning/sa   keso   ] su   [background pig-bakal-an kaso babayi su   tindahan].

            NOM/GEN/DAT cheese NOM   LV-buy   GEN woman NOM store

        Intended: 'It's (the) cheese that the woman bought at the store.'

    From these examples, we see that local clefting can only target the pivot, the argument in nominative case and cross-referenced by voice morphology on the verb, and thus follows the claimed pivot-only restriction

on Ā-extraction (5c). As noted above, DP *wh*-questions are also formed using clefts and therefore follow the extraction restriction observed in (10–12).

3.2 Topics

Next, we turn to topics in Bikol. We use the term "topic" to refer to DP arguments in pre-verbal position without an exhaustive focus interpretation; see note 9 below. Topics can be formed in two different ways in Bikol: topicalization and hanging topic left dislocation (HTLD). We will argue that topicalization involves movement, whereas hanging topics are base-generated high. In the following examples, topics — and corresponding pronouns, if any — are in bold.

The examples in (13) involve topicalization of their pivots. (13a) has topicalized an agent pivot from an AV clause and (13b) has topicalized a theme pivot from a PV clause. Topicalization is associated with no intonational break and cannot be resumed by corresponding pronouns.

(13)   **Pivot topicalization:**

    a.   **Su   babayi** nag-kaon ning keso.

        NOM woman AV-eat   GEN cheese

    b.   **Su   keso**   k<in>aon kaso babayi.

        NOM cheese PV-eat   GEN woman

        'The woman ate (the) cheese.'

In contrast, hanging topics are followed by an obligatory intonational break and have a corresponding post-verbal pronoun. Consider the examples in (14) below. In (14a), the agent pivot 'woman' is topicalized from an AV clause, followed by an intonational break — indicated by a # — with a corresponding post-verbal nominative pronoun =*siya* which encliticizes to the verb. In (14b), the theme pivot 'cheese' is topicalized from a PV clause, with a following pause and corresponding full pronoun.

(14)   **Pivot HTLD, with prosodic break and pronoun:**

    a.   **Su   babayi # nag-kaon *(=siya)   ning keso.

        NOM woman   AV-eat   NOM.3sg GEN cheese

        'The woman$_i$, she$_i$ ate cheese.'

    b.   **Su   keso**   # k<in>aon kaso babayi *(**ito**).

        NOM cheese   PV-eat   GEN woman   NOM.DEM

        'The cheese$_i$, the woman ate it$_i$.'

The obligatoriness of the intonational break and corresponding pronoun are generally one-to-one. (An exception is discussed in footnote 24 below.) Throughout this paper, we will give English translations with canonical word order for Bikol examples with topicalization, as in (13) above, whereas we give English translations with hanging topics with corresponding pronouns for Bikol HTLD, as in (14). We have chosen to do this to highlight the presence or absence of the corresponding pronoun in the Bikol sentences through their English translations. We should however reiterate that we are making no claims regarding the discourse status of these two constructions that we call "topics" here and, in particular, we make no claim that the information-structural properties of these Bikol sentences match those of their English translations.[9]

With these basic descriptions of the two forms of topics in place, we now consider which arguments can be targeted for topicalization and HTLD. Examples (13) and (14) above showed that both topicalization and HTLD can target pivots. Topicalization can additionally target the non-pivot agent of Non-Actor Voice clauses. This is observed in the PV example (15), and we will see the same generalization extend to LV below.

(15)   **Non-pivot agent topicalization:**

Su   **babayi** k<in>aon su   keso.

NOM woman PV-eat   NOM cheese

i.   'The woman ate the cheese.'

ii. * 'The cheese ate the woman.'

Note that the topic in (15) must be in nominative case, even though the corresponding post-verbal position is a genitive case position. Example (16) retains the original genitive case marker on the topic *babayi* 'woman' in (15), resulting in ungrammaticality. Recall that multiple post-verbal arguments cannot be in nominative case; see (9) above.

---

[9] As Shibatani (1991: 116) notes, "the status of the initial topic construction in Philippine languages does not seem to be well established yet." Unfortunately we are unaware of substantial progress on this important question, three decades later, especially as it concerns Bikol in particular. Here we follow Shibatani and others in referring to these preverbal phrases as "topics." In grammaticality and felicity judgment tasks, our speakers accept sentences with topicalization (13a,b) and HTLD (14a,b) in the same contexts as their corresponding post-verbal pivot clauses and comment that the choice is optional. The determination of the discourse effects of these topic constructions — if any — awaits further work, potentially requiring corpus and/or experimental methods.

(16)   **Ungrammatical topicalization of non-pivot agent with original case marker:**   cf (15)

   ***Kaso babayi** k\<in>aon su   keso.

   GEN   woman PV-eat   NOM cheese

On the surface, topicalizing a non-pivot agent as in (15) results in a string with two nominative phrases: the pre-verbal topic 'woman' and the post-verbal 'cheese.' However, (15) is unambiguous in its interpretation: the post-verbal nominative phrase is unambiguously the pivot of this PV clause and therefore the verb's theme, whereas the pre-verbal nominative topic is unambiguously the non-pivot agent.[10] Our proposal below will account for this restriction.

Although non-pivot agents of transitive verbs can be topicalized, non-pivot themes cannot. This is illustrated in (17) below, which attempts to topicalize the non-pivot theme *keso* 'cheese' from an AV clause. The sentence is ungrammatical with *keso* in nominative or its original genitive case.

(17)   **Topicalization of non-pivot theme is ungrammatical:**

   ***Su/ning/sa**   **keso**   nag-kaon su   babayi.

   NOM/GEN/DAT cheese AV-eat   NOM woman

   Intended: 'The woman ate (the) cheese.'

The generalization that topicalization can target pivots and non-pivot agents (15) but not non-pivot themes (17) also extends to clauses with transitive verbs in additional voices. Consider the options for topicalization from a Locative Voice (LV) clause in (18). Interestingly, in LV, where both the agent and theme are non-pivot arguments as determined by the choice of voice morphology, we continue to observe an asymmetry: topicalization can target the non-pivot agent (18b), again resulting in a structure with two nominative phrases, but cannot target the non-pivot theme (18c). The locative pivot can also naturally be topicalized, as in (18a).

---

[10] This is not simply an effect of animacy. For example, example (2b) in section 1, which is structurally parallel to (15), is similarly unambiguous.

(18) **Topicalization from a LV clause:**

a. <u>Grammatical locative pivot topicalization:</u>

Su tindahan pig-bakal-an kaso babayi ning keso.

NOM store   LV-buy   GEN woman GEN cheese

b. <u>Grammatical non-pivot agent topicalization:</u>

*Su/*kaso babayi* pig-bakal-an su   tindahan ning keso.

NOM/GEN woman LV-buy   NOM store   GEN cheese

c. <u>Ungrammatical non-pivot theme topicalization:</u>

* **Su/ning keso** pig-bakal-an kaso babayi su   tindahan.

NOM/GEN cheese LV-buy   GEN woman NOM store

'The woman bought cheese at the store.'

The one exception to the unavailability of non-pivot theme topicalization comes from unaccusative predicates. Example (19a) is a baseline that shows that the non-pivot theme of unaccusative 'fall' receives genitive case in Locative Voice; example (19b) shows that this argument can be topicalized, again requiring nominative case in place of its post-verbal genitive.

(19) **Topicalization of unaccusative theme from a LV clause:**

a. <u>Baseline unaccusative in Locative Voice:</u>

Pig-takdag-an ni   Maria su   halong.

LV-fall   GEN Maria NOM house

b. <u>Grammatical non-pivot theme topicalization:</u>

*Si/*ni*   **Maria** pig-takdag-an su   halong.

NOM/GEN Maria LV-fall   NOM house

'Maria fell into the house.' (for example from an airplane)

In summary, topicalization — which we will argue below to involve movement — does not follow a pivot-only restriction, unlike clefting in (10–12). Topicalization can target non-pivot agents and non-pivot unaccusative themes, although non-pivot themes are otherwise inaccessible. Topicalization is thus freer than clefting but not unrestricted.

Next we turn to hanging topic left dislocation (HTLD). We saw in example (14) above that HTLD can target pivots. In addition, HTLD can target non-pivot agents as well as non-pivot themes of transitive clauses

as in (20–21) below. These examples each correspond to the topicalization examples in (15) and (17) above, where we saw that non-pivot agents but not non-pivot themes can be topicalized.

(20) **Non-pivot agent HTLD from PV and LV:**

a. *Su* **babayi #** k<in>aon **=niya** su keso.
   NOM woman PV-eat GEN.3sg NOM cheese
   'The woman$_i$, she$_i$ ate the cheese.'

b. *Su* **babayi #** pig-bakal-an **=niya** su tindahan ning keso.
   NOM woman LV-buy GEN.3sg NOM store GEN cheese
   'The woman$_i$, she$_i$ bought cheese at the store.'

(21) **Non-pivot theme HTLD from AV and LV:**

a. *Su* **keso #** nag-kaon su babayi **kaito**.
   NOM cheese AV-eat NOM woman DAT.DEM
   'The cheese$_i$, the woman ate it$_i$.'

b. *Su* **keso #** pig-bakal-an kaso babayi **kaito** su tindahan.
   NOM cheese LV-buy GEN woman DAT.DEM NOM store
   'The cheese$_i$, the woman bought it$_i$ at the store.'

In these examples of non-pivot HTLD (20–21), the topics themselves are in nominative case, even though their corresponding pronouns are in genitive or dative case.[11] Like (15) and (18b) above, the resulting string has two nominative phrases, but each is unambiguous in its interpretation. The pre-verbal hanging topic must correspond to the post-verbal pronoun.

We conclude that there is no restriction on the DP arguments that can be targeted by HTLD. We will argue that this is because HTLD does not involve movement, in contrast to topicalization.

3.3 Summary

In this section, we presented data on clefting and two types of topics from local clauses in Bikol. Local clefting obeys the pivot-only extraction restriction. Topicalization can target pivots and non-pivot agents, as well as non-pivot unaccusative themes. Hanging topic left dislocation can target any core argument, including

---

[11] We gloss the pronouns in (21) as dative as they are specific non-pivot themes, which appear in dative case; see e.g. (8). However, *kaito* is both a genitive and dative form, as noted in footnote 5.

non-pivot themes of transitive verbs. These possibilities are summarized in (22) below for arguments of notionally transitive verbs.

(22) **Possible DP targets for local dependencies from transitive clauses:**

	pivot	non-pivot agent	non-pivot theme
clefting	✓ (10a, 11a, 12a)	* (11b, 12b)	* (10b, 12c)
topicalization	✓ (13, 18a)	✓ (15, 18b)	* (17, 18c)
hanging topic left dislocation	✓ (14)	✓ (20a,b)	✓ (21a,b)

All non-pivot topics involve an apparent mismatch in case marking: the pre-verbal topic is in nominative case, instead of the genitive or dative case of its corresponding post-verbal gap or pronoun. In the next section, we present our analysis for Bikol voice and case, as well as the specific analyses for clefting, topicalization, and HTLD, with additional supporting data.

4 Proposal

In this section we present our analysis for the patterns of voice, case, and local dependencies in Bikol introduced in the previous section. A key point that we account for is the ability of topicalization to target non-pivot agents as well as pivots, but not non-pivot themes of transitives, in contrast to clefting which is strictly pivot-only and hanging topic left dislocation (HTLD) which is unrestricted. To preview our account, we will propose that topicalization is a movement construction that involves probing for a discourse feature [TOP], restricted only by Phase Impenetrability, whereas clefting involves a probe that seeks a focused DP that must be the closest DP to the probe (see also Erlewine, 2018; Branan and Erlewine, to appearb). The pivot and non-pivot agent are the only DPs at the *v*P phase edge, and so there is no way to target non-pivot themes, even with a [TOP] probe, except where *v*P does not introduce an impenetrable barrier as with unaccusatives. In contrast, HTLD is a non-movement construction, unrestricted by Phase Impenetrability. We will also discuss the determination of morphological case in Bikol, explaining the appearance of multiple nominative phrases in some topic constructions.

Our proposal is presented in three parts. Section 4.1 presents our proposal for case and voice in Bikol. We present our analysis for the two topic constructions in section 4.2 and for clefts in section 4.3. We then further motivate our analysis of the Bikol clause periphery from patterns of multiple topics in section 4.4. Note that all dependencies in this section will be local, accounting for the patterns presented in section 3 above. We then discuss long-distance clefting in section 5.

4.1 Voice and case in Bikol

We begin by presenting our framework for the voice system and morphological case in Bikol. For the voice system, we follow the shared insights of widely-adopted and influential phase-based approaches to voice systems in Philippine languages, drawing especially on the work of Aldridge (2004, 2008) and Rackowski and Richards (2005). Under such approaches, the pivot DP is distinguished by being the highest DP in vP — the lower phase of the clause — in a (outer) specifier of vP. Agents are base-generated in Spec,vP. In Actor Voice (AV) clauses, there is no movement to the edge of the vP phase; the agent pivot is base-generated as the only specifier of vP and remains the highest DP in the vP; see (23a). In Non-Actor Voice (NAV) clauses, a non-agent DP is moved to the outer specifier of vP, above the agent DP (23b).[12] Specifiers of vP are illustrated on the left in trees, but this does not reflect their word order, as we discussed below.

(23)   **The vP phase in AV and NAV clauses:**

    a.   <u>Actor Voice:</u>

vP

DP
agent
pivot

v

...

    b.   <u>Non-Actor Voices:</u>

vP

DP
pivot

DP
agent

v

... t ...
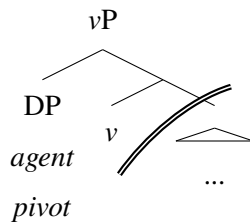
vP is a phase and therefore material within the complement of the phase head v will be inaccessible for syntactic operations from above (Phase Impenetrability), except in the case of unaccusative verbs (Chomsky, 2000). In (23), this domain of impenetrability is illustrated with a double line. T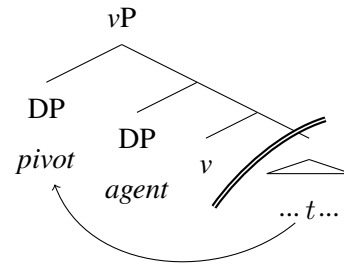his approach predicts a basic asymmetry between AV and NAV clauses: in AV clauses, the vP phase edge has only one DP that is accessible for syntactic operations from above, whereas in NAV clauses, there are two. In the following subsection, we will propose that this is precisely what allows for topicalization to target only pivots (in AV and NAV clauses) and non-pivot agents (in NAV) of transitive clauses; these are the only DP constituents of the lower phase that can move out. See also Erlewine and Levin 2021 for a recent, additional argument for precisely this organization of the vP phase edge, based on the inventory of clitic pronouns in Philippine-type voice system languages, including Bikol. In the case of unaccusatives, the complement of v is accessible for

---

[12] Complement CPs may also function as the pivot and move to the outer specifier of v. This will be important in the analysis of long-distance dependencies, discussed in section 5.

probing and therefore non-pivot themes can be topicalized, as we have seen above.[13]

Voice morphology is the realization of the head *v*, which the lexical verb V head-moves to. Aldridge 2004 and Rackowski and Richards 2005 differ in the precise mechanisms that relate the realization of voice morphology to movement of the pivot DP in NAV clauses. However, both of these approaches agree on the basic geometry for the *v*P phase edge in AV vs NAV clauses, reviewed in (23) above. We adopt this common proposal here. NAV clauses involve movement of the pivot DP to an outer specifier of *v*P whereas AV clauses involve no such movement, leaving the agent to be the highest DP in the phase and the only DP at in the *v*P phase edge.[14, 15]

Post-verbal word order in Bikol is free, except for a requirement that complement clauses be rightmost. We adopt the proposal from Erlewine, Levin, and Van Urk 2020 (also in Erlewine 2018 for Toba Batak) that all linearizations of *v*P with the verbal complex (*v*+V) as the leftmost constituent can be generated. See also Fowlie 2013 and Branan to appear for similar proposals for Tagalog.

Next we turn to our proposal for morphological case determination in Bikol. Following Marantz 1991,

---

[13] Editor Julie Anne Legate points out that we predict that both internal arguments of a double unaccusative would be accessible for topicalization. At this point, we have not been able to identify appropriate predicates in the language to test this prediction.

[14] Pivots of peripheral voices must either be generated as DPs via an applicative functional head without a lower source for structural case licensing, as Aldridge and Rackowski & Richards assume, or may originate as a prepositional/oblique object that is then severed from its preposition via preposition-incorporation (Baker, 1988); the latter approach is discussed in Guilfoyle et al. 1992: fn. 7 and Kroeger 1990 for Tagalog and Bodily 2003 for Cebuano. For our current purposes, we leave this analytic choice open and concentrate on PV examples as illustrative of NAV syntax more generally.

[15] As pointed out by our editor and a reviewer, and recently discussed extensively in Hsieh 2020 ch. 7, non-DP extraction (in Tagalog) generally proceeds differently from DP extractions, and in particular does not interact with voice choice. The same appears to be true in Bikol; see for example (49) in section 4.3 below for an instance of pre-verbal PP fronting which does not interact with voice and which behaves differently from pre-verbal DPs.

As Hsieh points out, such behaviors are at first glance problematic for phase-based accounts of voice system languages and their extraction restrictions. In particular, prominent prior works such as Aldridge 2004 and Rackowski and Richards 2005 do not discuss any mechanism for moving non-DP constituents to the phase edge in order to feed their extraction from above. However, a recent theoretical proposal in Van Urk and Richards 2015 on the Nilotic language Dinka rectifies the issue. (Van Urk 2015 and Erlewine, Levin, and Van Urk 2015, 2017 observe that Dinka exhibits all of the characteristic properties of Austronesian voice systems in (5) above.) Van Urk and Richards (sec. 4) propose that *v* in Dinka is associated with two features, one Case/D-related and another Ā-related; when the two can be satisfied by a single Ā-DP goal, the probes must target that one goal together, but when there is an Ā-non-DP distinct from the pivot DP, both that non-DP and the pivot can be moved to separate specifiers of *v*P, simultaneously. Voice morphology choice is only sensitive to the movement of DPs. This mechanism or similar allows for the principled movement of non-DPs through the lower phase edge without interaction with voice morphology.

we propose that morphological case in Bikol may be structurally assigned or realized with context-sensitive defaults.[16] In particular, we propose that default case in the *v*P phase is genitive and default case in the CP phase is nominative. The idea of genitive as a default case within some structures in Austronesian languages is developed by Donohue and Donohue (2010), Chen (2018), and Erlewine, Levin, and Van Urk (2020). An anonymous reviewer further points out that this choice of genitive as default within the lower domain of the clause may have a diachronic origin, due to the lower portion of clauses in many Austronesian languages having its origin in nominal structures. In addition, nominative can be assigned structurally by T, via Agree, as in Aldridge's (2004) analysis of absolutive in "T-type" languages.

The derivation of AV and NAV clauses as well as the determination of morphological case will be illustrated below. We begin with the transitive AV clause derivation in (24). Following the voice system proposal above in (23), the agent is base-generated in Spec,*v*P and no other argument is moved to the *v*P phase edge. We propose that T bears [PROBE:D] which assigns structural nominative case to its target.[17] As the agent is the highest DP in the *v*P — and, in this case, the only one accessible by Phase Impenetrability — [PROBE:D] on T necessarily targets the agent pivot, which receives nominative case.

(24)   **Actor Voice clause derivation:**
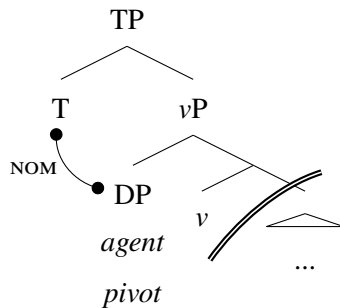
Any DP that is realized in the *v*P phase and lacks structural case-marking will receive default genitive case (Erlewine, Levin, and Van Urk, 2020). This accounts for the genitive case on non-pivot themes in AV clauses. In addition, as noted above, specific non-pivot themes receive dative case through a separate process

---

[16] Marantz (p. 247) proposes that "unmarked case may be sensitive to the syntactic environment; for example, in a language GEN may be the unmarked case for NPs inside NPs (or DPs) while NOM may be the unmarked case inside IPs." See Levin 2015: ch. 6 for recent discussion of Marantz 1991. Relevant for us is Levin's discussion which concludes that "[Marantz's] categories of unmarked and default case must be collapsed" (p. 212). We use the term "default" here. See also Baker 2015 and Chen 2018 for recent work that further develops this idea of context-sensitive default case.

[17] For explicitness, we use the notation [PROBE:F] to represent a probe for the feature [F]. This avoids the more common [uF] notation, which different linguists take to stand for "unvalued" or "uninterpretable."

(see footnote 6) and therefore will not receive default genitive. The surface form computed for an AV clause with a non-specific theme is presented in (25). Recall that the linear order of constituents in the *v*P is subject to scrambling, with the only constraint being that the verbal complex be leftmost.

(25) **An Actor Voice clause at PF:**

"AV-V NOM=DP$_{ag/pivot}$ GEN=DP$_{th}$" *or* "AV-V GEN=DP$_{th}$ NOM=DP$_{ag/pivot}$"

Next we turn to the derivation of transitive Non-Actor Voice clauses. This is illustrated with the tree in (26). As we introduced above, in NAV clauses, a non-agent DP moves above the agent to an outer specifier of *v*P. [PROBE:D] on T will find the highest DP, which is the pivot, and assign it structural nominative case. The non-pivot agent has not received structural case, so it will receive default genitive as it is in the *v*P.

(26) **Non-Actor Voice clause derivation:**
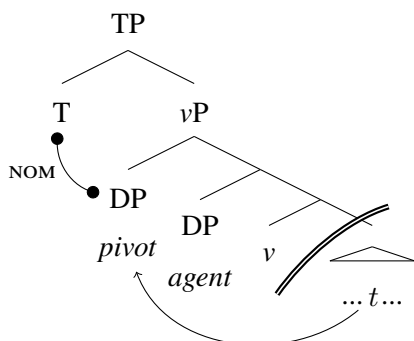
(27) **A Non-Actor Voice clause at PF:**

"NAV-V NOM=DP$_{pivot}$ GEN=DP$_{ag}$" *or* "NAV-V GEN=DP$_{ag}$ NOM=DP$_{pivot}$"

It's worth highlighting that the *v*P phase boundary is relevant here in two distinct senses. For purposes of probing and movement, the complement of the phase head *v* constitutes a distinct domain, inaccessible for higher probing (Phase Impenetrability), except where the verb is unaccusative. This boundary is indicated by the double line in the trees above. However, for purposes of linearization (scrambling) and default case calculation, it is the entire *v*P maximal projection, including its specifiers, that behaves as a unit in all clauses. Unless moved higher, specifiers of *v*P are linearized post-verbally and subject to scrambling together with all other *v*P-internal constituents. Non-pivot agents receive default genitive case, just as (non-specific) non-pivot themes do. We suggest that this distinction correlates with the timing of the relevant operations: probing is a narrow-syntactic operation and is sensitive to the double line (Phase Impenetrability), whereas linearization and default case determination takes place post-syntax, at PF, where the entire *v*P behaves as a single unit.

Finally, we discuss the calculation of morphological case for a constituent that moves out of *v*P. First consider the movement of pivots. Pivots receive structural nominative and will retain this structural case

19

when moved. However, the situation is more complicated when a non-pivot DP moves. Due to Phase Impenetrability, this is only possible with non-pivot agents, as illustrated in (28), or with non-pivot themes of unaccusatives.

(28)   **Movement of a non-pivot agent:**
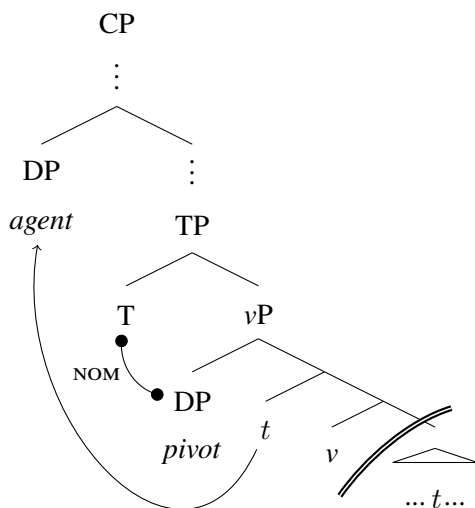
We propose that any DP without structural case that is pronounced in the CP phase will receive default nominative. Non-pivot agents have no source of structural case, so their morphological case realization will depend on the phase in which they are pronounced. If the agent stays within the *v*P phase, it appears with default genitive. But if the agent moves out into the CP phase, as in (28), it will appear in nominative case. The PF realization of a structure as in (28) is sketched in (29). This description also extends to moved non-pivot themes of unaccusatives, as in (19b).

(29)   **A moved non-pivot agent at PF:**
       "**NOM**=DP$_{agent}$ ... [$_{v}$P NAV-V   NOM=DP$_{pivot}$]"

There are therefore two sources of surface nominative case in our proposal: structural nominative via Agree with T and default nominative in the CP phase. In (29), the post-verbal pivot DP bears structural nominative whereas the pre-verbal non-pivot agent bears default nominative by virtue of its position in the CP phase. As noted by Schütze (2001), identity between structural nominative and a default case in a higher domain of the clause (e.g. on topics) is cross-linguistically common. [18]

---

[18] An alternative approach would be to claim that all nominative is the result of default nominative in CP, and that Agree with T has the effect of making a DP target behave as if it is in the CP for the purposes of default case calculation. That is, even in cases

A consequence of this proposal is that, despite there being multiple sources of nominative case, only one DP can bear nominative case and appear in a post-verbal position. This is the pivot DP which receives structural nominative from T while in the vP and therefore linearized post-verbally. For any other DP to bear nominative case, it must move out of the vP into the CP phase and therefore be in a pre-verbal linear position. This explains the impossibility of multiple post-verbal nominatives, as illustrated in (9) above.

The analysis for Bikol voice and case presented in this section derives the surface morphosyntax for basic AV and NAV clauses in Bikol that we saw in section 2. In addition, two features of this approach will be important for the analysis of Bikol topics and clefts, which we turn to in the following sections. First, the new proposal that nominals in the CP receive default nominative will be important for deriving the case marking observed on topics. Second, two DPs are at the vP phase edge in NAV clauses — the pivot and the non-pivot agent — whereas only the agent pivot is at the phase edge in AV clauses. While this is a feature of previous phase-based accounts for voice system syntax in Rackowski 2002, Aldridge 2004, and Rackowski and Richards 2005, its consequences have not been fully discussed in previous work (except recently in Erlewine and Levin 2021). This organization of the vP phase edge will be crucial for explaining the differing extraction restrictions on clefting vs topicalization in Bikol.

4.2 Topicalization and hanging topic left dislocation

We showed in section 3.2 that there are two topic constructions in Bikol: topicalization, which involves a gap and no prosodic break, and hanging topic left dislocation (HTLD), which has a corresponding pronoun and a prosodic break. Topicalization can target non-pivot agents as well as pivot DPs, but not non-pivot themes (except from unaccusatives), whereas HTLD can target any DP argument. In this section we present our analysis for these facts.

We propose two functional heads in the clause periphery, which we simply label Top_2 and Top_1, with Top_2 c-commanding Top_1. In Rizzi 1997 terms, these can be thought of as heads in a split CP. This organization is illustrated schematically in (30):

---

where the pivot stays in the vP, the pivot will receive the CP's default nominative, due to its Agree relationship with the CP phase's T head. This alternative proposal has the conceptual advantage of describing all surface forms of nominative case as due to the same mechanism — default case in the CP — but with a new proposal regarding the effect of Agree with T. We will leave the full consideration of this alternative description for future work.

(30) **Topic layers in the Bikol clause periphery:**

```
                    CP
                   /  \
                  C    Top₂P
                      /    \
              hanging      /  \
              topic,    Top₂   Top₁P
              (HTLD)         /     \
                        moved      /  \
                        topic   Top₁   TP
                     (topicalization)
```
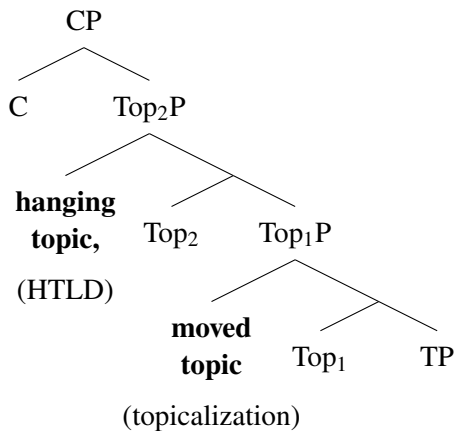
Topicalization is due to Top_1: [PROBE:TOP] on Top_1 fronts any [TOP] goal it finds to Spec,Top_1P. Top_2 generates hanging topics: a DP is base-generated in Spec,Top_2P and binds a pronoun in its scope.[19] Any constituent in Spec,Top_2P is followed by a prosodic break. In section 4.4, we present data from multiple topicalization that supports the higher position for hanging topics.

The claim that topicalization involves movement while HTLD involves base-generation and binding is supported by differences in island-sensitivity (Ross, 1967). Examples (31–32) below show that topicalization but not HTLD is sensitive to islands, as diagnosed by examples with attempted topic dependencies into an adjunct island (a) or relative clause island (b).

(31) **Topicalization is island-sensitive:**

    a. *Su   babayi pig-uran [$_{island}$ bagu pig-hiling ni   Andrew].

        NOM woman PV-rain   before PV-see   GEN Andrew

        Intended: 'It rained [$_{island}$ before Andrew saw the woman].'

    b. *Su/sa   eskwela na-hiling =ko   su   lalaki [$_{island}$ na nag-gadan].

        NOM/DAT student PV-see   GEN.1sg NOM man   that AV-kill

        Intended: 'I saw the man [$_{island}$ that killed the student].'

---

[19] We can think of the Top_2 head itself as having the semantics of a binder, as in Büring 2005 or similar to Constant 2014's functional head for contrastive topics.

(32) **HTLD is island-insensitive:**

a. Su   babayi # pig-uran [island bagu pig-hiling =siya   ni   Andrew].
   NOM woman   PV-rain   before PV-see   NOM.3sg GEN Andrew

   'The woman_i, it rained [island before Andrew saw her_i].'

b. Su   eskwela # na-hiling =ko   su   lalaki [island na nag-gadan sainya].
   NOM student   PV-see   GEN.1sg NOM man   that AV-kill   DAT.3sg

   'The student_i, I saw the man [island that killed them_i].'

Further evidence for this movement / non-movement contrast comes from the interpretation of verb-argument idiom chunks (see e.g. Marantz, 1984). Here we use two idioms for 'mumbling' and 'being a coward':

(33) **Baselines with idiomatic interpretations:**

a. Piga-kaon ni   Andrew su   tataramon.
   PV-eat   GEN Andrew NOM words/language

   'Andrew mumbles / doesn't speak clearly.'   (literally 'Andrew eats words/language')

b. Ma-luya su   buot   =ko.
   AV-weak NOM feelings GEN.1sg

   'I am a coward / lack courage.'   (literally 'my feelings are weak')

Topicalization retains these idiomatic interpretations, in (34), but HTLD does not, leaving only their literal interpretations available, in (35). This is explained by the topics in (34) being generated together with their predicates and then subsequently moved, whereas the hanging topics in (35) are base-generated high and thus never in a local relationship with their predicates.

(34) **Topicalization retains idiomatic interpretation:**

a. Su   tataramon   piga-kaon ni   Andrew.
   NOM words/language PV-eat   GEN Andrew

   'Andrew mumbles / doesn't speak clearly.'

b. Su   buot   =ko   ma-luya.
   NOM feelings GEN.1sg AV-weak

   'I am a coward / lack courage.'

(35)   **HTLD blocks idiomatic interpretations:**

    a. #Su   tataramon   # piga-kaon ito   ni Andrew.

        NOM words/language   PV-eat   NOM.DEM GEN Andrew

        'Words/language$_i$, Andrew eats them/it$_i$.'

    b. #Su   buot   =ko   # ma-luya ito.

        NOM feelings GEN.1sg   AV-weak NOM.DEM

        'My feelings$_i$, they$_i$ are weak.'

We now turn to the explanation for the possible targets of topicalization. Probing is subject to Phase Impenetrability (Chomsky, 2000); therefore, [PROBE:TOP] on Top$_1$ cannot probe into the complement of *v* and attract a matching goal. In AV clauses, this means that only the pivot agent can be topicalized (23a). In NAV transitive clauses, two DPs are potentially accessible for probing: the pivot and the non-pivot agent, which are both specifiers of *v*P; see (23b). Non-pivot themes are not accessible for topicalization because of Phase Impenetrability, except where the verb is unaccusative and therefore the complement of *v* is accessible for probing from above. This accounts for the patterns of topicalization documented in section 3: pivots and non-pivot agents are the only DPs that can be topicalized from transitive clauses, with non-pivot themes of unaccusatives also topicalizable. [PROBE:TOP] will find the closest accessible target with the [TOP] feature. In cases of non-pivot topicalization, the non-pivot bears a [TOP] feature but the higher pivot does not. Because the pivot does not bear the feature that the probe seeks, it does not intervene for the topicalization of the non-pivot.

DPs without structural case in Spec,Top$_1$P or Spec,Top$_2$P will be realized with default nominative case; see (29) above. This explains the appearance of nominative case on non-pivot topics, as in (36), which correspond to a post-verbal gap or genitive or dative pronoun. In addition, pivots which receive structural nominative and are moved to Spec,Top$_1$P also appear in nominative case.

(36)   **Nominative case on non-pivot topics:**

    a. **Su** **babayi** k\<in\>aon su   keso.   = (15)

        NOM woman PV-eat   NOM cheese

        'The woman ate the cheese.'

    b. **Su** **babayi** # k\<in\>aon =niya   su   keso.   = (20a)

        NOM woman   PV-eat   GEN.3sg NOM cheese

        'The woman$_i$, she$_i$ ate the cheese.'

c. **Su keso** # nag-kaon su babayi kaito. = (21b)
   NOM cheese   AV-eat   NOM woman DAT.DEM

   'The cheese$_i$, the woman ate it$_i$.'

Because only pivots can be both post-verbal and in nominative case, the interpretation of such examples with multiple nominatives is unambiguous, as noted above.

4.3 Clefts

Recall that, unlike the possibilities for local topics, local clefting is strictly pivot-oriented:

(37) **Grammatical local pivot cleft:** = (11a)

   [focus Su lalaki] su [background g<in>adan kaso eskwela].
       NOM man NOM   PV-kill   GEN student

   'It's the man that the student killed.'

(38) **Ungrammatical local non-pivot agent cleft:** = (11b)

   *[focus Su eskwela] su [background g<in>adan su lalaki].
       NOM student NOM   PV-kill NOM man

   Intended: 'It's the student that killed the man.'

We propose that clefting is the result of attraction by a probe that seeks a target with both [FOC] and [D] features that must be the closest DP to the probe. Probing of this form is discussed and motivated in Erlewine 2018 and Branan and Erlewine to appearb, and following these works we notate this probe [PROBE:FOC+D]. Ā-probing that is restricted to the closest DP has been a component of some analyses of syntactic ergativity as in Aldridge 2004, but is also well attested in non-ergative languages, as shown in Branan and Erlewine to appearb. This probe will match the highest DP if it bears [FOC], but will not match a non-[FOC] highest DP and also cannot probe past it for a better match.[20] As discussed in Branan and Erlewine to appeara, such a probe can be described in Deal's (2015; to appear) interaction–satisfaction model of probing as [INT:FOC+D, SAT:D].

---

[20] Our proposal could account for the extraction restriction on clefting if we instead describe clefting as involving [PROBE:D]; this in turn would be functionally equivalent to Aldridge's (2004; 2017) proposal where, in her terms, extraction involves a [uφ] probe. However, our description here of clefting involving [PROBE:FOC+D] that must target the closest DP better reflects the need of clefting to target a focused constituent, which must be independently enforced in some way under a [PROBE:D] analysis.

Pre-verbal DP exhaustive focus constructions — which we call "clefts" here — in Bikol and many related Philippine languages invite two types of analyses, schematized in (39) below. The first is a biclausal pseudocleft structure, in (39a), where the focused DP constitutes a predicate nominal in a null copular clause, and the cleft background is formally a headless relative clause. The pre-verbal nominative marker is then explained as the predicted nominative case marker for the subject of the copular clause. The second is a monoclausal structure where a functional head Foc extracts the DP focus out of the background clause. On this latter approach, we might associate the observed pre-verbal nominative marker *su* with the Foc head itself, with the motivation for this surface form being that the monoclausal structure (39b) historically derived from the biclausal pseudocleft structure (39a).

(39)   **Two modes of description for Philippine clefts:**

  a.   [predicate   DP$_{focus}$]   *su*$_{NOM}$ [background V ... ___ ...]

  b.   [FocP   DP$_{focus}$ *su*$_{Foc}$ [background V ... ___ ...]]

Our core proposal for clefting in Bikol and its extraction restriction is in principle compatible with either analytical approach, as long as the probe involved is [PROBE:FOC+D] restricted to the closest DP and the background is not a full CP, as we discuss. Here for concreteness we adopt the latter, monoclausal description as in (39b), for two reasons. First, our preliminary look at patterns of relativization in Bikol suggests that it is not strictly pivot-oriented, suggesting that the extraction restriction of clefts cannot be attributed to that of relativization.[21] Second, the structure of clefts, both above and below the focused DP, can be neatly modeled by positing a clause type that varies minimally from that in (30) in having the focus-attracting Foc head in place of Top$_1$, as illustrated in (40) below.[22]

---

[21] For instance, (i) and (ii) below are attested non-pivot agent and non-pivot theme relatives, respectively. We leave a complete investigation of patterns and mechanisms of relativization in Bikol for future work.

(i)   su   lalaki [RC na   pig-tama-an su   babayi]   (ii)   su   lalaki [RC na   nag-tama si   Maria]
       NOM   man   that LV-hit   NOM   woman   NOM   man   that AV-hit   NOM   Maria
       'the man that hit at the woman'   'the man that Maria hit'

(40)   **The Bikol clause periphery for clefts:**

```
              CP
            /    \
           C    Top₂P
               /    \
       hanging      /  \
       topic,   Top₂  FocP
       (HTLD)       /   \
                 cleft   /  \
                 focus  Foc  TP
                         |
                         su
```
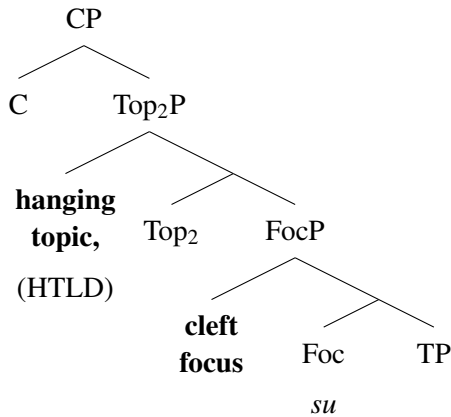
Below we will show that the cleft background clause — the sister of Foc, i.e. TP in (40) — has a structurally reduced left periphery as compared to matrix and embedded complement clauses, as is cross-linguistically common (see e.g. Belletti, 2012), and in particular cannot include the Top_1 and Top_2 projections introduced in the previous section. A hanging topic can however be hosted above a cleft focus, as reflected by the Top_2 head in (40). Finally, we note that under our analysis, the *su* marker in clefts is synchronically the realization of the Foc head (40) rather than a case marker, we will continue to gloss it as NOM throughout.

We first show that the cleft background cannot include a hanging topic nor movement-derived topic DP. As local clefting is limited to the pivot DP and topicalization can target the pivot or non-pivot agent of transitive verbs, the most plausible configuration to test would be with a pivot focus and non-pivot agent topic, as in (41) below. Example (41) is ungrammatical, with or without the pause and corresponding pronoun to make the topic a hanging topic.

(41)   **Cleft background cannot include a topic:**

*[focus Su   keso] su   [background **si**   **Andrew** (#) k<in>aon (**=niya**)   ].
      NOM cheese NOM   NOM Andrew   PV-eat   GEN.3sg

Intended: 'It's cheese that Andrew (, he) ate.'

---

[22] Scholars since at least Kroeger (1991) have observed that second-position clitics do not climb to fronted pre-verbal DP foci, but do climb to fronted non-DPs. Aldridge (2002, 2004) in particular takes such facts to form an argument that DP clefts are biclausal, as in (39a). (But see also Kaufman 2010: 188 fn 4 for critical discussion of this type of argument.) Parallel facts hold in Bikol; see (48) below. However, we note that clitics do not climb to pre-verbal topics (both movement-derived topics and hanging topics) in Bikol, and thus the domain of clitic climbing in both topic constructions as in (30) and in clefts as in (40) can be described uniformly as TP under our proposal. Evidence from clitic climbing thus does not motivate a biclausal analysis of DP clefts any more so than it does for DP topic constructions.

Additional, independent evidence for the cleft background not itself being a full clause comes from high adverbials. Consider the speaker-oriented modifier 'unfortunately' in (42). (43) shows that it can appear before or after a topicalized pivot.

(42)   **'Unfortunately' in pre-verbal and post-verbal positions:**

{Sa   kamalasan} nag-inom {sa   kamalasan} ning hilo   su   babayi.
   DAT bad.luck   AV-drink DAT bad.luck   GEN poison NOM woman

'Unfortunately the woman drank poison.'

(43)   **'Unfortunately' available before or after movement-derived topic:**

{Sa   kamalasan} su   babayi {sa   kamalasan} nag-inom ning hilo.
   DAT bad.luck   NOM woman DAT bad.luck   AV-drink GEN poison

'Unfortunately the woman drank poison.'

'Unfortunately' cannot appear at the left edge of a cleft background, although it is allowed post-verbally, as in (44). The availability of 'unfortunately' post-verbally shows that the adverb is semantically compatible with being in the cleft background, and thus that its ungrammaticality at the left edge of the cleft background is due to a syntactic restriction. The adverb can also appear above the cleft focus as in (45). These patterns further support our proposal that cleft backgrounds are not full clauses.

(44)   **'Unfortunately' cannot appear at the left edge of the cleft background:**

[focus Su   babayi] su   [background {*sa   kamalasan} nag-inom {sa   kamalasan} ning hilo   ].
        NOM woman NOM   DAT bad.luck   AV-drink DAT bad.luck   GEN poison

'It's the woman that unfortunately drank poison.'

(45)   **'Unfortunately' can appear above the cleft focus:**

Sa   kamalasan [focus su   babayi] su   [background nag-inom ning hilo   ].
   DAT bad.luck   NOM woman NOM   AV-drink GEN poison

'Unfortunately it's the woman that drank poison.'

In contrast, it is possible to have a hanging topic above the cleft focus, as in (46). In both examples in (46), the cleft focus is the pivot, according with the generalization that local clefting is limited to the pivot, with the hanging topic corresponding to a non-pivot core argument.

(46)   **Hanging topic can precede the cleft focus:**

a.   **Su   babayi** # [focus su   keso   ] su   [background k<in>aon **=niya**   ].

        NOM woman   NOM cheese NOM   PV-eat   GEN.3sg

        'The woman_i, it's the cheese that she_i ate.'

b.   **Su   keso**   # [focus su   babayi] su   [background nag-kaon **kaito**   ].

        NOM cheese   NOM woman NOM   AV-eat   DAT.DEM

        'The cheese_i, it's the woman that ate it_i.'

This supports our proposal in (40), which includes a higher Top$_2$ head to host hanging topics above cleft foci.

We propose that clefting involves movement of the focused constituent, predicting clefting to be island-sensitive. This is demonstrated with the adjunct island and relative clause island data in (47). The island-sensitivity of clefting here patterns with topicalization in (31) but stands in contrast to the island-insensitivity of HTLD in (32).

(47)   **Clefting is island-sensitive:**

a.   *[focus Su   babayi] su   [background pig-uran [island bagu   pig-hiling ni   Andrew]].

        NOM woman NOM   PV-rain   before PV-see   GEN Andrew

        Intended: 'It's the woman that it rained [island before Andrew saw *t*].'

b.   *[focus Su   eskwela] su   [background na-hiling =ko   su   lalaki [island na   nag-gadan]].

        NOM student   NOM   PV-see   GEN.1sg NOM man   that AV-kill

        Intended: 'It's the student that I saw the man [island that killed *t*].'

Recall too that local clefts must have a post-verbal gap in the background clause, corresponding to the focus, further supporting their derivation via movement. (See examples (10–11) above.) This detail will become important in section 5, where we will see that long-distance clefts may have a resumptive pronoun in place of a gap.

Finally, our description of clefts as triggered by [PROBE:FOC+D] predicts that non-DP categories such as PPs cannot undergo clefting. Although PPs can also be focus-fronted, as in (48), it is clear that this does not involve the same structure as the DP focus clefts described here. First, a pre-verbal focused PP cannot be followed by the *su* marker which is required for DP focus clefts. Second, second-position clitics such as the pronoun =*ako* in (48) can be hosted on these focused pre-verbal PPs, but second position clitics do not climb up to the DP focus of clefts, as has also been described for Tagalog (Kroeger, 1991: 123–125); see also footnote 22.

(48) **Pre-verbal PP focus is not a cleft:**

Sa eskwelahan **=ako** (**su*) nag-basa ning libro, bako sa halong.

DAT school NOM.1sg NOM AV-read GEN book NEG DAT home

'It's at school that I read books, not at home.'

These facts suggest that non-DP fronting constructions are markedly different from DP fronting constructions; see also footnote 23. We leave their detailed study for future work and refer readers to Hsieh 2020 for recent discussion of such structures in Tagalog.

A related prediction that our theory of clefting makes, as noted by an anonymous reviewer, is that clefting via [PROBE:FOC+D] should be able to skip intervening non-DP categories. This prediction is also borne out. Consider example (49), based on the cleft in (37) above, which shows that it is possible to have a locative PP before the verb within a cleft background:

(49) **Clefting skips intervening PPs:**

[focus Su lalaki] su [background **sa halong** g<in>adan kaso eskwela].

NOM man NOM at home PV-kill GEN student

'It's the man that the student killed at home.'

Recall from (41) above that the edge of the cleft background cannot host DP topics. Therefore the pivot 'man' must have moved from a post-verbal position within the background clause, within vP. The pre-verbal PP is necessarily outside of vP. The grammaticality of (49) thus shows that clefting is unaffected by intervening PPs, as is predicted by our proposal that clefts are derived using [PROBE:FOC+D] on Foc.[23]

We note that both aspects of our proposal for the structure of clefts — an Ā-probe that necessarily targets the closest DP (Erlewine, 2018; Branan and Erlewine, to appearb) and the lack of topic projections within the cleft background (41) — are necessary to derive the strict pivot-only restriction on local clefting, reflected in (37–38). If we were to use a simple [PROBE:FOC] that is not limited to the closest DP, we would predict that a non-[FOC] pivot DP could be skipped, allowing the cleft to attract a [FOC]-bearing non-pivot agent or unaccusative non-pivot theme instead, just as we proposed that topicalization involves [PROBE:TOP]. At the same time, if the cleft background contained Top$_1$ or Top$_2$, a topic could be built first, making a non-pivot

---

[23] Example (49) also shows that, despite the unavailability of high adverbials and DP topics in cleft backgrounds (41–44), there is space to host locative PPs. This again reinforces the idea that the positions available for pre-verbal DPs and pre-verbal PPs in Bikol fundamentally differ, as we also noted above in relation to focused DPs vs focused PPs. See also footnote 15 for a suggestion regarding the analysis of non-DP movements.

argument the highest DP within the background clause. Subsequent clefting with [PROBE:FOC+D] would be predicted to be able to attract that non-pivot argument, fed by topicalization or HTLD within the background clause. Therefore, to derive the pivot-only restriction on local clefts, the background clause must not be a full CP, as we argued above.

4.4 Multiple topic constructions

We now return to the two topic constructions in Bikol and show that the analysis presented above in section 4.2 is further supported by examples with multiple pre-verbal topics. This section serves to present these positive predictions of our account, as well as to establish baseline behaviors that will become important for our discussion of patterns of long-distance clefting in section 5.3 below.

We first consider the two grammatical PV examples in (50). Both topics in (50) are in nominative case, as is independently predicted for each topic construction.

(50) **PV clauses with multiple topics:**

    a.   Si   Pedro # su   babayi   g<in>adan (=niya).

        NOM Pedro   NOM woman   PV-kill   GEN.3sg

        i.   'Pedro$_i$, he$_i$ killed the woman.'

        ii. * 'Pedro$_i$, the woman killed him$_i$.'

    b.   Si   Pedro # su   babayi   g<in>adan =siya.

        NOM Pedro   NOM woman   PV-kill   NOM.3sg

        i. * 'Pedro$_i$, he$_i$ killed the woman.'

        ii.   'Pedro$_i$, the woman killed him$_i$.'

Both examples in (50) are PV clauses with two pre-verbal DPs — Pedro and *babayi* 'woman' — but they differ in their interpretation, depending on the choice of post-verbal pronoun. In (50a), Pedro is the agent, corresponding to the post-verbal genitive pronoun, while *babayi* is the pivot theme.[24] In (50b), Pedro is the pivot theme, corresponding to the post-verbal nominative pronoun, while *babayi* 'woman' is the agent. Both examples are unambiguous in their interpretation.

The generalization is as follows. In these sequences of two topics, the first topic is a hanging topic, with a prosodic break and corresponding post-verbal pronoun, whereas the second topic is the result of

---

[24] But note that the genitive clitic pronoun is optional in this example. We nonetheless analyze Pedro in (50a) as the result of HTLD, as it must be followed by its characteristic prosodic break, but with the pronunciation of its bound pronoun being optional. At this point we are unable to describe the precise conditions that govern this optionality, and leave this question open for future work.

topicalization. Example (51) below shows that it is not possible to add a prosodic break after the second topic, with or without a break after the first topic, and regardless of the choice of post-verbal pronoun.

(51) **No prosodic break possible after the second DP:**

    *Si   Pedro (#) su   babayi **#** g<in>adan (=niya/=siya).

    NOM Pedro   NOM woman   PV-kill   GEN.3sg/NOM.3sg

This data in (50–51) supports our proposal that topics with a prosodic break and corresponding pronoun (hanging topics in Spec,Top$_2$P) are structurally higher than topics with no break and no corresponding pronoun (movement-derived topics in Spec,Top$_1$P).

Let's consider the derivation of each of these PV multiple topic examples in (50). We first consider the derivation of (50a). Here there is a hanging topic binding an agent pronoun and a topicalized theme pivot. We therefore begin by constructing a PV clause with the full DP 'woman' with a [TOP] feature as the theme and a pronoun as the agent. Following movement of the pivot theme to an outer specifier of the vP, we result in a vP organized as in (52):

(52) **vP for (50a):**
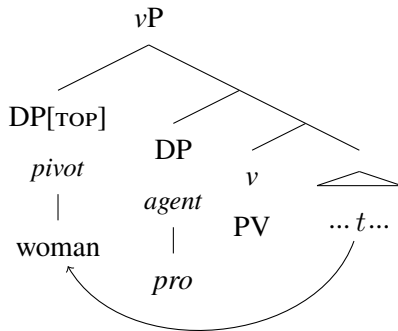
The rest of the clausal spine is built following the hierarchy in (30), beginning with the merger of T. [PROBE:D] on T will Agree with the closest DP, assigning *babayi* 'woman' nominative case. The agent pronoun is in vP so it receives default genitive case. Top$_1$ is merged and its [PROBE:TOP] fronts the pivot DP 'woman' to Spec,Top$_1$P. Top$_2$ is then merged in and takes Pedro as its specifier, which binds the lower agent pronoun. The clause is complete once we merge the C head to form the root CP. The resulting hierarchical structure is as in (53a) below, together with its final linearized structure in (53b).

(53) **Final structure for (50a):**

    a.   [$_{CP}$ [$_{Top2P}$ Pedro$_i$ Top$_2$ [$_{Top1P}$ woman Top$_1$ [$_{TP}$ T [$_{vP}$ $t_{th/pivot}$ [$pro_{ag,i}$ [v ...

b. "NOM=DP$_i$ # NOM=DP$_{th/pivot}$ PV-V =GEN.3sg$_{ag,i}$"

⇒ "si Pedro$_i$, su babayi ginadan=niya$_i$"

'Pedro$_i$, he$_i$ killed the woman.'

Both topics are realized in nominative case: *babayi* 'woman' bears structural nominative from T whereas Pedro receives default nominative in the CP. The hanging topic in Spec,Top$_2$P is followed by a prosodic break. The post-verbal pronoun is genitive and thus appears in the *=niya* form. This results in the correct surface form attested in (50a/53b), and also derives the correct, unambiguous interpretation for this string.

Next we turn to the derivation of example (50b). This example is superficially similar to (50a) but with a post-verbal nominative pronoun in place of the genitive pronoun in (50a), resulting in a markedly different interpretation, 'Pedro, the woman killed him.' We begin by building a PV *v*P with a pronoun theme pivot moving to its outer specifier, above the [TOP]-marked agent DP *babayi*.

(54) **vP for (50b):**
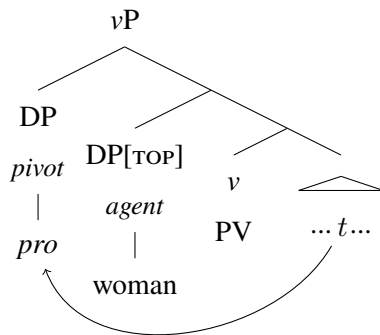
We now build the higher phase. T is merged and [PROBE:D] assigns nominative case to the pivot pronoun, which is the closest DP goal. Next, Top$_1$ is merged and its [PROBE:TOP] moves the agent to Spec,Top$_1$P. The Top$_2$ head is merged with its specifier, Pedro, which binds the theme pivot pronoun. After merging in C, we yield the structure in (55):

(55) **Final structure for (50b):**

a. [$_{CP}$ [$_{Top2P}$ Pedro$_i$ Top$_2$ [$_{Top1P}$ woman Top$_1$ [$_{TP}$ T [$_{vP}$ *pro* $_{th/pivot,i}$ [t_{ag} [v ...

b. "NOM=DP$_i$ # NOM=DP$_{ag}$ PV-V =NOM.3sg$_{th/pivot,i}$"

⇒ "si Pedro$_i$, su babayi ginadan=siya$_i$"

'Pedro$_i$, the woman killed him$_i$.'

Both topic DPs receive default nominative case because they are in the CP phase. The lower pronoun is also nominative as it is the pivot and so received structural nominative, resulting in the post-verbal clitic form =*siya*. This results in the correct surface form in (50b/55b), with the correct interpretation.

So far we've looked at multiple topics in a PV clause. Under our proposal both the pivot and non-pivot agent in a transitive NAV clause are at the *v*P phase edge and thus accessible for topicalization, and both arguments can be targeted for HTLD as well. This allowed for the two minimally contrasting examples in (50) above which are both grammatical but with differing interpretations. But now consider multiple topics in an AV transitive clause. Here we observe an asymmetry: Example (56a) is grammatical with its post-verbal dative pronoun, whereas (56b) is ungrammatical with its post-verbal nominative pronoun.

(56)   **AV clauses with multiple topics:**

    a.   Su   eskwela **#** su   lalaki nag-gadan sainya.

        NOM student   NOM man   AV-kill   DAT.3sg

        i.   'The student$_i$, the man killed them$_i$.'

        ii. * 'The student$_i$, they$_i$ killed the man.'

    b. *Su   eskwela **#** su   lalaki nag-gadan =siya.

        NOM student   NOM man   AV-kill   NOM.3sg

        Intended: 'The student$_i$, they$_i$ killed the man.'

This asymmetry is predicted by our account. Following our proposal and the discussion of the PV examples in (50) above, the outer, hanging topic *eskwela* 'student' in (56) must bind the post-verbal pronoun, with the inner topic *lalaki* 'man' being moved from its base position. In an AV clause, only the agent pivot is at the *v*P phase edge and thus available for topicalization. In contrast, HTLD is not similarly limited as it does not involve movement. This together explains the grammaticality of example (56a). Example (56b) is ungrammatical because the non-pivot theme *lalaki* 'man' would have to be moved from within the lower phase, in violation of Phase Impenetrability. This asymmetry observed in AV clauses with multiple topics in (56) thus further supports both our analysis for the difference between topicalization and HTLD as well as our proposal for the syntax of the *v*P phase edge in AV and NAV clauses, following Rackowski 2002, Aldridge 2004, Rackowski and Richards 2005, and Erlewine and Levin 2021.

4.5   Summary

In this section we presented our proposal for Bikol clause structure, morphological case, topics, and clefts. Concentrating on the salient difference between the two movement operations of topicalization and clefting,

we proposed a locality-based account for the differing extraction restrictions.

Our analysis builds on common Minimalist assumptions regarding the locality of syntactic operations. In particular, movement is subject to Phase Impenetrability and is triggered by a probe that must target its closest goal (Chomsky, 2000, 2001 and many others). The pivot-only restriction on clefting is due to a probe that seeks a target with both [FOC] and [D] features and which must be the closest DP, as motivated further in Branan and Erlewine to appearb. Topicalization instead involves [PROBE:TOP] which can skip a non-[TOP] pivot DP to attract a non-pivot argument. Phase Impenetrability explains the inability of topicalizing non-pivot themes, which are not at the *v*P phase edge, except in unaccusatives where *v*P does not form a barrier for probing.

Finally, we note that under our view, nothing about clefting is inherently linked to pivot-hood. As proposed in section 4.3 above, background clauses of a local cleft are simply structured so that the pivot is necessarily the closest DP to the probe. We predict that if there is a strategy for making a non-pivot DP closer to the cleft's probe, clefting would target this non-pivot DP instead. We will see that this is the case in the next section, where we consider long-distance clefts.

5 Long-distance clefts and the Austronesian extraction restriction

In this section, we take a closer look at the nature of the famed Austronesian pivot-only extraction restriction. We have seen that, in Bikol, this restriction is obeyed by local clefting but not by local topicalization or HTLD, so our approach will be to further study clefting in Bikol. At first glance, there are several different ways to characterize this type of extraction restriction:

(57)   **Three characterizations of the classic Austronesian extraction restriction:**

    i.   <u>Pivot-only</u>: Only arguments cross-referenced by voice ("pivots") can be extracted.

    ii.   <u>Nominative-only</u>: Only nominative arguments can be extracted.

    iii.   <u>Locality-based</u>: Only the structurally highest argument can be extracted.

The challenge is to distinguish between these three different descriptions. Every clause has only one pivot, which is in nominative case. Assuming that a topic cannot be formed first (see section 4.3 above), every clause also only has one nominative argument, which is the pivot. And assuming the basic proposal for the hierarchical structure of voice system languages (section 4.1 above), the highest argument in every clause will be the pivot, in nominative case. Therefore, in basic examples of local clefting, these three descriptions in (57) are extensionally equivalent: In the background clause of local clefts, the pivot is the only nominative

argument, and is structurally highest. The study of local clefts alone does not allow us to determine the correct characterization for the extraction restriction.

For this reason, in this section we study long-distance clefting in Bikol. We begin in section 5.1 with some preliminary discussion of long-distance extraction in voice system languages. The core data on long-distance clefting will be presented in section 5.2. Unlike in local clefts, long-distance clefting can target embedded non-pivot agents as well as embedded pivots, which forms an argument against the "pivot-only" characterization of clefting in (57i). We propose that, in such examples, embedded topicalization takes place first and feeds clefting. We support this approach, in section 5.3, with additional data from the interaction of long-distance clefting and embedded topics. In the end, we will also be able to tease apart the "nominative-only" (57ii) and locality-based (57iii) approaches, solidifying our argument that the Austronesian extraction restriction exemplified by Bikol clefting must be described in terms of hierarchical structural configurations and the locality of syntactic operations.

5.1   Background: Voice systems and long-distance extraction

Just as Ā-extraction from local clauses is limited in languages with Austronesian-type voice systems, long-distance extraction is also similarly constrained. Descriptively, extraction out of an embedded clause in Bikol requires that the embedded clause itself be the pivot of the higher clause. In other words, long-distance Ā-movement is always subextraction from a clausal pivot. This pattern has been well-documented in Tagalog since Kroeger 1991: ch. 7, and is also a major point of discussion in Rackowski and Richards 2005.

The examples in (58) illustrate the grammatical long-distance clefting of embedded pivots. The pivots of the embedded clauses are the theme Andrew in (58a), where the embedded clause is PV, and the agent 'man' in (58b), where the embedded clause is AV.

(58)   **Long-distance clefting is grammatical across a PV verb:**

    a.   Si   Andrew su   [$_{TP}$ pig-balita ning radyo [$_{CP}$ na g<in>adan kaso lalaki.
        NOM Andrew NOM   PV-report GEN radio   that PV-kill   GEN man
        'It's Andrew that the radio reported that the man killed *t*.'

    b.   Su   lalaki su   [$_{TP}$ pig-balita ning radyo [$_{CP}$ na nag-gadan ki   Andrew.
        NOM man   NOM   PV-report GEN radio   that AV-kill   DAT Andrew
        'It's the man that the radio reported that *t* killed Andrew.'

Notice that in both cases the higher verb 'report' is in PV, with its agent 'radio' in genitive case as expected. We can think of the complement clause 'that the man killed Andrew' as the pivot of the verb 'report' in PV,

although CPs do not exhibit morphological case marking. The embedded clause's pivot is then subextracted to yield the grammatical cleft in (58).

The higher verb must be PV for this long-distance extraction to take place. Example (59) below minimally contrasts from (58a), with the higher 'report' clause now in AV, and the result is ungrammatical. The pivot of this higher clause is the agent 'radio,' instead of the complement clause.

(59)   **Long-distance clefting is ungrammatical across an AV matrix clause:**

   *Si   Andrew su   [TP **nag**-balita su   radyo [CP na   g<in>adan kaso lalaki.

   NOM Andrew NOM   AV-report NOM radio   that PV-kill   GEN man

   Intended: 'It's Andrew that the radio reported that the man killed *t*.'

Under the probe-driven conception of movement adopted here, what is important for our purposes is that the highest DP within the embedded CP count as the "closest" for the cleft's [PROBE:FOC+D] in (58), instead of the agent DP of the verb 'report.' Different approaches could be taken, but for concreteness here we briefly present and follow the analysis of long-distance extraction from Rackowski and Richards' study of Tagalog. In grammatical cases of long-distance extraction as in (58), the complement CP itself moves to an outer Spec,*v*P above any agent DP. The verb is in the PV form, correlating with this movement of the theme to Spec,*v*P. This structure is illustrated in (30). Recall that *v*P will be linearized with the verbal complex leftmost, explaining the final word order. CPs are generally rightmost, due either to extraposition or their relative weight.

(60)   **An embedded CP in Spec,*v*P (PV), as in (58):**

⇒ "*v*+V DP_{agent} CP"
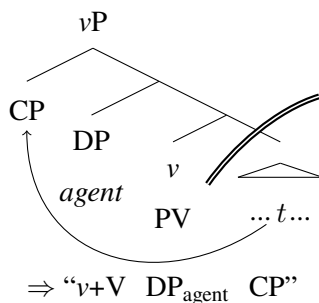
Movement of the CP here "smuggles" the target DP above the higher clause's agent DP. In particular, Rackowski & Richards propose that the relationship between *v* and CP makes the CP transparent for probing from above.[25] The cleft's [PROBE:FOC+D] will thus search into the pivot CP, matching with the highest DP goal within. As discussed in Branan and Erlewine (to appeara), such "smuggling" derivations (see also Collins 2005; Belletti and Collins 2021) may be made possible by a depth-first search procedure (or similar,

as in Chow 2022) that looks into the contents of accessible specifiers before proceeding down the spine, or by first targeting the CP as a partial match (as per the intuition in the aforementioned works) and then searching further within it.

In contrast, if the higher verb is in AV as in (59), the complement CP will not move to Spec,*v*P. Due to Phase Impenetrability, it is impossible to probe into the CP that is inside the lower VP.

(61)   **An embedded CP in the complement of *v* (AV), as in (59):**

*v*P
DP
agent
pivot
v
AV
⋮
... CP ...

The licit and illicit patterns of probing from above for a goal in the lower phase of a transitive clause are summarized in (62). In simple cases of probing for a local goal, the goal must be in Spec,*v*P to be accessible for probing from above (62a–b) due to Phase Impenetrability, making pivots and non-pivot agents uniquely visible for probing from above. In cases where the goal is embedded within a CP, that CP itself must move to Spec,*v*P to escape Phase Impenetrability (62c–d) and to be made transparent for probing (see footnote 25).

(62)   **Patterns of probing:**

a.   *PROBE ... [$_{v}$P v_{AV} [$_{VP}$ ... GOAL

b.   PROBE ... [$_{v}$P GOAL v_{PV} [$_{VP}$ ...

c.   *PROBE ... [$_{v}$P v_{AV} [$_{VP}$ ... [$_{CP}$ ... GOAL ...]   (59/61)

d.   PROBE ... [$_{v}$P [$_{CP}$ ... GOAL ...] v_{PV} [$_{VP}$ ...   (58/60)

For these reasons, in all subsequent examples of long-distance clefting, the higher verb will be in PV. Such examples become ungrammatical with a different choice of voice marking, as in (59) above.

---

[25] Independent evidence from Zulu (Halpert, 2012, 2016: ch. 6, Halpert, 2019), Dinka (Van Urk and Richards, 2015), and Hungarian (Den Dikken, 2018: ch. 4) support the idea that an Agree relationship with a CP is necessary for probing into it. Van Urk and Richards (2015) argue that, in the Nilotic language Dinka, the embedded CP additionally moves to Spec,*v*P to feed extraction out of the CP.

5.2 Long-distance clefting

Long-distance clefting in Bikol differs from local clefting in two ways. First, long-distance clefting can involve a gap or a resumptive pronoun in the embedded clause, whereas local clefts must involve a gap. Second, gapped long-distance clefts can target non-pivot agents as well as pivots, unlike local clefts which must target the pivot. We begin with discussion of the latter property. We have seen in example (58) above that embedded pivots can be clefted long-distance. Example (63), repeated from (4) above, shows that embedded non-pivot agents can also be clefted long-distance.

(63) **Long-distance cleft of embedded non-pivot agent:**   = (4)

Su   eskwela su   [TP pig-balita ning radyo [CP na g<in>adan su   lalaki.
NOM student NOM   PV-report GEN radio   that PV-kill   NOM man

'It's the student that the radio reported *t* killed the man.'

We propose that long-distance clefting of non-pivot agents as in (63) involves a first step of embedded topicalization, followed by long-distance clefting. First, we note that topicalization can take place within embedded clauses, moving a non-pivot agent to the embedded CP clause edge. Just as in topicalization in local matrix clauses, the non-pivot agent topic *eskwela* 'student' appears in nominative case in (64).

(64) **Embedded topicalization of a non-pivot agent:**

Pig-balita ning radyo [CP na **su   eskwela** g<in>adan su   lalaki.
PV-report GEN radio   that NOM student PV-kill   NOM man

'The radio reported that the student killed the man.'

The embedded non-pivot agent *eskwela* 'student' is now the highest DP in the embedded CP in (64). If we cleft from (64), [PROBE:FOC+D] will search into the embedded CP, as the higher verb 'report' is PV, and attract the highest DP in the embedded clause, if focused. This allows for the successful derivation of the long-distance non-pivot agent cleft in (63).

Now recall that Bikol also has another way to form topics, HTLD, associated with a prosodic break and a corresponding pronoun. HTLD can also target embedded CP edges, as demonstrated with an embedded non-pivot agent hanging topic in (65). If the DP generated in this embedded hanging topic position bears [FOC], clefting using [PROBE:FOC+D] based on a structure as in (65) will yield a long-distance non-pivot agent cleft with a resumptive pronoun instead of a gap, which is indeed grammatical, in (66).

(65) **Embedded HTLD of a non-pivot agent:**

Pig-balita ning radyo [CP na **su   eskwela #** g<in>adan **=niya**   su   lalaki.

PV-report GEN radio   that NOM student   PV-kill   GEN.3sg NOM man

'The radio reported that, the student$_i$, they$_i$ killed the man.'

(66) **Long-distance non-pivot agent cleft with resumptive pronoun:**

Su   eskwela su   [TP pig-balita ning radyo [CP na   g<in>adan **=niya**   su   lalaki.

NOM student NOM   PV-report GEN radio   that PV-kill   GEN.3sg NOM man

'It's the student$_i$ that the radio reported that *they$_i$* killed the man.' = (63)

Note that the cleft focus in (66) is not followed by the prosodic break associated with hanging topics (65). This is, however, predicted by our account, where the prosodic break associated with HTLD is tied to the pronunciation of a constituent in Spec,Top$_2$P.

The derivations we propose for (63) and (66) involve [FOC]-targeting movement of an embedded movement-derived or base-generated topic. Higher, non-topic Ā-extraction targeting an embedded topic has been proposed for example in the derivation of long-distance clefts with resumptive pronouns in Berber (Shlonsky, 1987), long-distance relatives with resumptive pronouns in Moroccan Arabic (Elomari, 1998), and long-distance *wh*-questions with resumptives in colloquial French (Sportiche, 2018), as well as for long-distance relatives both with and without resumptives in Sportiche 2017.

In contrast to topicalization, HTLD can target all arguments, including non-pivot themes of transitive verbs. This predicts that an embedded non-pivot theme of a transitive verb can be clefted long-distance as long as it is fed by embedded HTLD, not topicalization, making the corresponding embedded pronoun obligatory. This is borne out in (67).

(67) **Long-distance non-pivot theme cleft requires resumptive pronoun:**

Su   eskwela su   [TP pig-balita ning radyo [CP na   nag-gadan su   lalaki *(**sainya**).

NOM student NOM   PV-report GEN radio   that AV-kill   NOM man   DAT.3sg

'It's the student$_i$ that the radio reported that the man killed *them$_i$*.'

We also predict that non-pivot themes of unaccusatives can be clefted long-distance, fed by embedded topicalization and thus leaving a gap. This is borne out, in example (68), based on example (19b) above.

(68) **Long-distance cleft of embedded non-pivot unaccusative theme:**

Si	Maria	su	[TP pig-balita	ning	radyo	[CP na	pig-takdag-an	su	halong.
NOM	Maria	NOM	PV-report	GEN	radio	that	LV-fall	NOM	house

'It's Maria that the radio reported *t* fell into the house.'

The availability of hosting both movement-derived topics and hanging topics at embedded clause edges reflects the fact that these embedded complement clauses are full CPs. This is in contrast to the background clause of clefts, which we argued in section 4.3 to be TP which for instance cannot host the high adjunct 'unfortunately.' In contrast, 'unfortunately' is available at the edge of embedded clauses in long-distance clefts, again reflecting their full CP size:

(69) **'Unfortunately' at edge of embedded clause, based on (63):**

Su	eskwela	su	[TP pig-balita	ning	radyo	[CP na	**sa**	**kamasalan**	g<in>adan	su	lalaki.
NOM	student	NOM	PV-report	GEN	radio	that	DAT	bad.luck	PV-kill	NOM	man

'It's the student that the radio reported unfortunately *t* killed the man.'

In this section we've concentrated on the possibility of topicalization or HTLD feeding clefting as a means of clefting embedded non-pivot arguments, but the same approach can also yield long-distance clefts of an embedded pivot DP. As predicted by this approach, long-distance pivot clefts as in (58) can also involve resumptive pronouns, which reflects embedded HTLD followed by clefting.[26]

(70) **Long-distance pivot cleft with resumptive pronoun, reflecting embedded HTLD:**

Si	Andrew	su	[TP pig-balita	ning	radyo	[CP na	g<in>adan	**=siya**	kaso	lalaki.
NOM	Andrew	NOM	PV-report	GEN	radio	that	PV-kill	NOM.3sg	GEN	man

'It's Andrew_i that the radio reported that the man killed *him_i*.' = (58)

The patterns of possible long-distance clefting of the arguments of an embedded transitive verb, with an embedded gap or resumptive pronoun, are summarized in (71) below, together with the possibilities for different local dependencies (22) repeated from section 3 above. As noted above, local and long-distance clefting differ in two ways: long-distance clefting can have a resumptive pronoun, while local clefting cannot, and long-distance clefting can target a greater range of possible DP arguments, also dependent upon the presence or absence of a pronoun.

---

[26] In the case of long-distance pivot clefting with a gap, as in (58) above, we hypothesize that its derivation may or may not involve a first step of embedded topicalization of the pivot. These derivations may be effectively indistinguishable.

(71) **Possible DP targets for local and long-distance dependencies from a transitive clause:**

	pivot	non-pivot agent	non-pivot theme
local clefting (gap)	✓	*	*
local topicalization (gap)	✓	✓	*
local HTLD (pronoun)	✓	✓	✓
long-distance clefting (gap)	✓ (58)	✓ (63)	* (67)
long-distance clefting (pronoun)	✓ (70)	✓ (66)	✓ (67)

Our proposal that embedded topicalization and HTLD can feed long-distance clefting predicts precisely this pattern in (71). Unlike the edge of a cleft background which is a TP (40), full CPs include the Top$_1$ and Top$_2$ projections, which can make a non-pivot the highest DP in the embedded CP, which can then be clefted.

(72) **Long-distance clefting fed by topicalization and HTLD:**

    a.   [$_{FocP}$ DP$_{focus}$ su_{Foc} [$_{TP}$ ... [$_{CP}$ na [$_{Top2}$ [$_{Top1}$ ___ [V ...   (58, 63, 68)

    b.   [$_{FocP}$ DP$_{focus}$ su_{Foc} [$_{TP}$ ... [$_{CP}$ na [$_{Top2}$ ___$_i$ [$_{Top1}$ [V pro_i ...   (66–67, 70)

Structures of the form in (72b), where material that is base-generated in an embedded clause and binds a lower pronoun is then moved higher, is a type of "mixed chain" in the terms of McCloskey 1979. Such mixed chain structures have been proposed in Irish (McCloskey, 1979, 2002: 195–197), Greek (Iatridou, 1995), Selayarese (Finer, 1997), Kaqchikel (Imanishi, 2019), and Dinka (Van Urk, 2017), in addition to the examples in Arabic, Berber, and French, from embedded topic positions mentioned above.

We have already argued in section 4 that topicalization involves movement, explaining why it is limited to pivots and a subset of non-pivot arguments and leaves a gap, whereas HTLD involves base-generation, explaining why it is not limited to particular arguments and involves a pronoun. Long-distance clefting with a gap can be derived by a first step of embedded topicalization, explaining why it is not strictly pivot-oriented, unlike local clefting. Clefting with a resumptive pronoun is only possible long-distance, because it is fed by embedded HTLD, and consequently can target any DP argument. This derives the pattern in (71).

We conclude with discussion of a potential conceptual complication to this proposal. Under some approaches to information structure, topic-hood and focus-hood are expected to be mutually incompatible. However, here we reiterate that we use the term "topic" (and the corresponding feature [TOP] as the trigger of movement topicalization) descriptively to refer to fronting that is not interpreted with exhaustive focus semantics. See footnote 9 above. In particular, our core findings here would be unaffected if the movement

here described as "topicalization" were instead thought of as a purely optional movement.[27]

5.3 Long-distance clefting and embedded topics

We have argued that long-distance clefting can involve a first step of embedded topicalization or HTLD. This approach then predicts non-trivial interactions between long-distance clefting and embedded topics. We will discuss such patterns in this section.

First, we observe that topicalization and HTLD can simultaneously target an embedded clause edge, just as they can simultaneously target the edge of a simplex clause, as we saw in section 4.4 above.[28] For ease of presentation, in this section we will use single and double underlines respectively for outer, base-generated hanging topics and inner, movement-derived topics, as well as their corresponding gaps. The two examples in (73) below are string-identical except for the choice of pronoun in the embedded clause and this correlates with their different interpretations.[29]

(73) **Multiple topicalization at an embedded clause edge:**

    a.   Pig-balita ning radyo [$_{CP}$ na   su lalaki # su eskwela g<in>adan $^?$(=niya).

            PV-report GEN radio   that NOM man   NOM student PV-kill   GEN.3sg

        i. * 'The radio reported that the man$_i$, the student killed him$_i$.'

        ii.   'The radio reported that the man$_i$, he$_i$ killed the student.'

    b.   Pig-balita ning radyo [$_{CP}$ na   su lalaki # su eskwela g<in>adan =siya.

            PV-report GEN radio   that NOM man   NOM student PV-kill   NOM.3sg

        i.   'The radio reported that the man$_i$, the student killed him$_i$.'

        ii. * 'The radio reported that the man$_i$, he$_i$ killed the student.'

In (73a), the hanging topic *lalaki* 'man' is interpreted as the non-pivot agent, whereas in (73b), it is interpreted as the pivot theme. Just as we established above in section 4.4 for unembedded multiple topic constructions,

---

[27] Following the common Minimalist assumption that all movement is feature-driven, this "optional" movement could be implemented by positing an optional feature [X] of no information-structural import, with the head we have described as "Top$_1$" in (30) bearing [PROBE:X] instead of [PROBE:TOP]. See Douglas 2018 for a recent proposal for topicalization via such a formal feature [X].

[28] In particular, recall that there are greater possibilities for multiple simultaneous topics in a Non-Actor Voice clause than in an Actor Voice clause. (See (50) vs (56) above.) We therefore use examples with embedded PV clauses here.

[29] The optionality of the genitive clitic pronoun in (73a) and (74a) below parallels its optionality in unembedded contexts as well, as we observed in section 4.4. The notation "$^?$(=niya)" reflects that the use of the pronoun =*niya* is strongly preferred in these examples, as we would expect of HTLD.

the generalization is that the post-verbal pronoun unambiguously corresponds to the higher, hanging topic. See (52–55) above for the derivation of these patterns, which also apply to the embedded clauses in (73).

The question now is what options are possible when we build clefts from these structures in (73). On the surface, the resulting clefts in (74) appear as long-distance clefts of *lalaki* 'man' with a single topic at the edge of the embedded clause. The two examples in (74) again differ only in the choice of pronoun after the embedded verb, and each example is unambiguous in its interpretation. Descriptively, the embedded resumptive pronoun corresponds to the fronted cleft focus.

(74) **Long-distance clefts with an embedded topic:**

    a. <u>Su   lalaki</u> su   [$_{TP}$ pig-balita ning radyo [$_{CP}$ na   <u>su   eskwela</u> g<in>adan $^?$(=niya).

        NOM man NOM   PV-report GEN radio   that NOM student PV-kill   GEN.3sg

      i. * 'It's the man$_i$ that the radio reported that the student killed him$_i$.'

      ii.   'It's <u>the man</u>$_i$ that the radio reported that he$_i$ killed <u>the student</u>.'

    b. <u>Su   lalaki</u> su   [$_{TP}$ pig-balita ning radyo [$_{CP}$ na   <u>su   eskwela</u> g<in>adan =siya.

        NOM man NOM   PV-report GEN radio   that NOM student PV-kill   NOM.3sg

      i.   'It's <u>the man</u>$_i$ that the radio reported that <u>the student</u> killed him$_i$.'

      ii. * 'It's the man$_i$ that the radio reported that he$_i$ killed the student.'

Notice that the interpretations of (74a,b) correspond one-to-one to the interpretations of examples (73a,b) above. That is, each example in (74) is unambiguously interpreted as a cleft of the embedded hanging topic 'man' from (73). We indicate this in (74) with corresponding gaps in the embedded hanging topic positions.

We can also be certain that the 'man' in the grammatical (74a,b) has indeed moved from the embedded clause as indicated, rather than being base-generated at the top. As we described in section 5.1 above, movement out of an embedded clause is only possible if the embedded clause functions as the pivot of the higher verb. Long-distance movement in (74a,b) was possible because the higher verb 'report' is in PV. If the higher clause is instead in AV, both (74a,b) become ungrammatical:

(75) **Variants of (74) with AV embedding verb are all ungrammatical:**

    * Su   lalaki su   [$_{TP}$ **nag**-balita su   radyo [$_{CP}$ na su   eskwela g<in>adan (=niya/=siya).

        NOM man NOM   AV-report NOM radio   that NOM student PV-kill   GEN.3sg/NOM.3sg

The unavailability of the (i) interpretation for the string in (74a) also teaches us that it is not possible to extract a post-verbal pivot across a pre-verbal topic. Consider a derivation where we begin with the embedded clause in (76a). If we were able to cleft the post-verbal pivot 'man' out of this embedded clause,

across the pre-verbal agent topic 'student,' we would predict the availability of the structure in (76b) as a long-distance theme pivot cleft. This result in (76b) is string-identical to (74a) and would be predicted to have the unattested (74ai) interpretation.

(76)   **Extraction of a post-verbal pivot across a topic:**

     a.   Embedded clause with topicalization:

          . . . [CP na   su   eskwela g<in>adan su   lalaki

                   that NOM student PV-kill   NOM man

          '. . . that the student killed the man.'

     b.   Hypothetical cleft of *lalaki* 'man' from (a), across the topic *eskwela* 'student':

         *Su   lalaki su   [TP pig-balita ning radyo [CP na   su   eskwela g<in>adan ___.

         NOM man NOM   PV-report GEN radio   that NOM student PV-kill

         Intended: 'It's the man that the radio reported that the student killed *t*.' = (74ai)

We return now to the derivation of clefts from the embedded multiple topic structures in (73). The examples in (74) showed that clefting of the outer, hanging topics from (73) is possible. What about clefting the inner, movement-derived topics from (73)? This would result in long-distance clefts with embedded hanging topics, marked by their characteristic prosodic gap, with the cleft foci corresponding to gaps in the embedded inner, movement-derived topic positions. These hypothetical structures with their predicted interpretations are given in (77). They are judged as ungrammatical.

(77)   **Long-distance clefting across an embedded hanging topic:**

     a.   *<u>Su   eskwela</u> su   [TP pig-balita ning radyo [CP na   <u>su   lalaki</u> # <u>___</u> g<in>adan (=niya).

         NOM student NOM   PV-report GEN radio   that NOM man   PV-kill   GEN.3sg

         Intended: 'It's <u>the student</u> that the radio reported that <u>the man</u>ᵢ, heᵢ killed <u>*t*</u>.'

     b.   *<u>Su   eskwela</u> su   [TP pig-balita ning radyo [CP na   <u>su   lalaki</u> # <u>___</u> g<in>adan =siya.

         NOM student NOM   PV-report GEN radio   that NOM man   PV-kill   NOM.3sg

         Intended: 'It's <u>the student</u> that the radio reported that <u>the man</u>ᵢ, <u>*t*</u> killed himᵢ.'

The structures in (74) vs (77) are presented schematically in (78) below. These patterns strengthen the argument that clefting necessarily attracts the DP that is highest and therefore structurally closest to the probe. From a structure with multiple embedded topics as in (73), it is only possible to cleft the higher, embedded hanging topic (74) and not possible to cleft the lower, embedded movement-derived topic (77).

(78) **Long-distance cleft from an embedded clause with two topics:**

    a.   [$_{FocP}$ DP$_{focus}$ su_{Foc} [$_{TP}$ ... [$_{CP}$ *na* [$_{Top2}$ $_i$ [$_{Top1}$ DP [V *pro*$_i$ ...   (74)

    b. *[$_{FocP}$ DP$_{focus}$ su_{Foc} [$_{TP}$ ... [$_{CP}$ *na* [$_{Top2}$ DP$_i$ [$_{Top1}$ _____ [V *pro*$_i$ ...   (77)

Similarly, from a clause with one pre-verbal DP topic and one post-verbal pivot, it is not possible to extract the pivot across the topic (76).[30]

    Moreover, recall from the previous section that topicalization can feed clefting in cases where there is no embedded hanging topic. This configuration is repeated here in (79) from (72a) above. The ungrammaticality of (78b) therefore cannot be attributed to a general immobility of inner, movement-derived topics.

(79) **Long-distance clefting can target a movement-derived topic:**   = (72a)

    [$_{FocP}$ DP$_{focus}$ su_{Foc} [$_{TP}$ ... [$_{CP}$ *na* [$_{Top2}$ [$_{Top1}$ _____ [V ...   (58, 63)

    Such data helps us to distinguish between the "nominative-only" and locality-based characterizations of the Austronesian extraction restriction in (57) above. Although all DPs that can be clefted are nominative in their lower positions (pivots or topics), being nominative is not a *sufficient* condition to be clefted. That is, the proper characterization of the restriction on clefting cannot be that *any* nominative phrase can be clefted. Even if being nominative is a prerequisite for clefting — see e.g. Deal 2017 on deriving extraction asymmetries through case-discriminating probing — only the highest nominative DP can be clefted. Because the structurally highest DP within any clause will necessarily be in nominative case (either a pivot which receives nominative case from T, or a topic which is realized with default nominative in CP), a restriction of the cleft's probe to nominative goals is unnecessary. The extraction restriction inevitably must refer to locality (the closest DP), and any characterization additionally referring to the nominative case of targets must be rejected on grounds of theoretical parsimony.

    A reviewer suggests that these long-distance clefting facts could also be addressed by taking cleft-formation to be topic-oriented, i.e. using [PROBE:TOP]. It is true that when there is an embedded topic, long-distance clefting targets the topic, and when there are multiple embedded topics, long-distance clefting targets the highest topic, as schematized in (78). In local clefting or with long-distance clefting without embedded topics, if we adopt the perspective that the pivot argument is necessarily a topic (see e.g. Chen,

---

[30] In contrast, recall that a post-verbal DP can be extracted across a pre-verbal PP; see example (49) above. This possibility is explained in our analysis by the fact that [PROBE:FOC+D] must attract the highest DP, but does not interact with intervening non-DP categories such as PPs.

2017), the reviewer suggests that we could describe clefting as uniformly attracting the closest topic. Aside from conceptual challenges to this approach, where clefts with exhaustive focus semantics are paired with a syntax that picks out a topic feature, cleft-formation via [PROBE:TOP] also fails to account for the fact that intervening non-DPs, which may potentially bear a [TOP] feature, must be skipped; see (49) above. Cleft-formation thus must specifically target the closest DP, which bears an Ā-feature that accords with the semantics associated with the construction; we therefore argue that it is best described as the result of [PROBE:FOC+D] that must target the closest DP. In the conclusion below, we also discuss additional challenges that the data here pose to such proposals that conflate the notions of pivot and topic.

This conclusion can also be translated into ergative hypothesis terms. As noted in section 2 above, many works describe Philippine voice system languages such as Bikol as exhibiting ergative/absolutive alignment. This includes Aldridge 2004, whose influential approach to the basic clause structure of voice systems we have adopted here. In brief, for these authors, what we have described here as nominative case is better described as absolutive and Ā-extraction in these languages exhibit "syntactic ergativity": in particular, an "absolutive-only" extraction restriction. If we were to adopt the ergative hypothesis as a mode of description, we would conclude that the syntactic ergativity observed in Austronesian voice system languages — evidenced in clefts in Bikol — in fact should not be described as an "absolutive-only" extraction restriction (pace e.g. Deal 2017). The appearance of this "absolutive-only" requirement on local clefts is due to the absolutive pivot argument being structurally highest in the cleft background clause. The source of this "syntactic ergativity" then is, again, best described as a locality-based effect: in particular, of Ā-probing for the closest DP, which is attested in both ergative and non-ergative languages (Branan and Erlewine, to appearb).

6 Conclusion

In this paper we've described and analyzed patterns of clefting and topic formation in Bikol, an Austronesian language of the central Philippines. Our analysis supports the view that the basic Austronesian "pivot-only" extraction restriction is best analyzed in terms of hierarchical configurations and the locality of syntactic operations. Ā-constructions that exhibit this "pivot-only" extraction restriction, such as cleft-formation in Bikol, involve Ā-probing that is restricted to targeting the closest DP (Branan and Erlewine, to appearb). This echoes Aldridge's (2004; 2017) earlier intuition that these constructions involve probing for $[\varphi]$ or $[D]$. In local clefts in Bikol, the pivot is necessarily the highest DP, but in long-distance clefting, topicalization or hanging topic left dislocation (HTLD) at the embedded clause edge can feed the cleft with a pivot or non-pivot argument as its closest DP target.

The behavior of Bikol long-distance clefts also forms an argument against case-discriminating approaches to the Austronesian extraction restriction. Based on the interactions of long-distance clefting and embedded topics studied in section 5.3, we conclude that consideration of syntactic locality is a necessary and sufficient condition for explaining the possible patterns for clefting in Bikol. Not only is there no preference for clefts to attract a "pivot," but it is both insufficient and unnecessary to describe clefting as subject to a case-discriminating (e.g. "nominative-only" or "absolutive-only") extraction restriction.

In contrast to clefting, topicalization in Bikol is not bound by the basic pivot-only restriction, even for local topics, but is also not completely unconstrained. The movement-derived construction of topicalization can target non-pivot agents as well as pivots, but not non-pivot themes except from unaccusatives. This is explained by the organization of the *v*P phase edge in Austronesian voice system languages: the agent is the only specifier of *v*P in Actor Voice, whereas in Non-Actor Voices, the pivot moves to an outer specifier of *v*P, resulting in two specifiers (Rackowski, 2002; Aldridge, 2004; Rackowski and Richards, 2005; Erlewine and Levin, 2021). *v*P is generally a phase for probing and extraction, with these specifiers of *v*P being the only possible targets for syntactic operations from above. Unaccusative *v*P however does not form a barrier for probing (Chomsky, 2000), allowing the topicalization of their non-pivot themes. This contrasts with the behavior of HTLD, which can target any DP argument, including non-pivot themes of transitive verbs. Evidence from island-sensitivity and idiom interpretation motivates the view that HTLD does not involve movement, unlike topicalization and clefting.

Our study also secondarily contributes to the theory of case determination, offering new evidence for the domain-sensitive approach to morphological case determination as in Baker 2015, where "default" case (Marantz, 1991) can take different forms in different domains of a single clause. In particular, we claim that genitive is the default case for DPs within the *v*P phase — as Erlewine, Levin, and Van Urk 2020 argue to be the case for many Austronesian languages and corresponding linearly to the post-verbal field — and nominative is the default case for DPs outside of *v*P in Bikol. Our evidence for this claim comes in particular from the behavior of movement-derived non-pivot topics, which are nominative but correspond to post-verbal genitive positions. The behavior of hanging topics, which are base-generated high, shows that nominative is the default case in the higher domain of the clause, which accords with the observation that nominative is often the default case for similar hanging topic constructions cross-linguistically (Schütze, 2001). Our analysis also offers a new, concrete theoretical approach to movement dependencies without case connectivity.

Finally, we conclude with a brief note on variation in the Austronesian extraction restriction(s), both within and between individual languages. We first note that some examples of non-pivot agent topics in other Philippine languages can be found in previous literature. The Tagalog example in (80) comes from

De Guzman 1995 and shows that non-pivot agents can be topicalized — with or without a corresponding pronoun — but non-pivot themes cannot.[31] Pizarro-Guevara 2020 has experimentally confirmed that Tagalog speakers accept non-pivot agent topics but not non-pivot theme topics, and that local non-pivot agent topics are much more acceptable than local non-pivot agent clefts. See also Ceña and Nolasco 2011 and Hsieh 2019: 528 fn 10; 2020: ch. 6; 2021 for more discussion of licit non-pivot agent extraction in Tagalog.

(80)   **Tagalog (De Guzman, 1995: 57):**[32]

    a.   **Ang nanay,** lulutu-in (**=niya**) ang isda sa kusina.

       NOM mother cook-PV GEN.3sg NOM fish DAT kitchen

       'The mother, (she) will cook the fish in the kitche.'

    b.   *****Ang isda,** mag-lulutu ang nanay sa kusina.

       NOM fish AV-cook NOM mother

       Intended: 'The fish, mother will cook (it).'

We have found similar examples of grammatical non-pivot agent topics but ungrammatical or unattested non-pivot theme topics in the Philippine languages of Hiligaynon (Mithun, 2019: 159), Limos Kalinga (Ferreirinho, 1993: 68–71), Kapampangan (Mirikitani, 1972: 154; Rowsell, 1983: 57–58), Pangasinan (Benton, 1971: 154), and Western Subanon (Blake, 2020). Reid (1978: 36) also presents parallel examples of this form from Bontok, Ilokano, Ivatan, and Tagalog. The same is observed of topicalization in Seediq, an Austronesian language of Taiwan (Atayalic), but where topics appear in a dedicated clause-final position instead of pre-verbally (Aldridge, 2004: 44-45); see also Erlewine 2014 for further discussion and Tsukida 2018: 320 for a similar example.

In each of these grammatical examples of non-pivot agent topicalization, the agent topic is in nominative case, resulting in a sentence with two nominative phrases, but the interpretation of the sentence is unambiguous, and the pronoun (if any) corresponding to the pre-verbal topic is in genitive case. These properties are exactly what we have observed in Bikol non-pivot agent topicalization. Discussing such examples, Shibatani (1988: 133) notes that "only those preposed subject [pivot] topics... are associated with

---

[31] Richards 2005 shows that non-pivot agents in Tagalog can also undergo another type of fronting called *ay*-topicalization, similarly resulting in structures with multiple *ang*-marked phrases, but non-pivot themes cannot. Miller 1988: 40–41 and Nagaya 2007: 366 give similar examples of non-pivot agent topics, also discussed in Kroeger 1991: 214–215, but the impossibility of non-pivot theme topics is not mentioned by these authors.

[32] Latrouite (2011b: 69) reports a variant of example (80b) with a post-verbal genitive demonstrative pronoun *nito* with the judgment mark ??.

focus [voice] marking in the verbs... the (pure) topics do not control focus [voice] marking." In other words, such examples motivate a clear distinction between the notion of pivot (his "subject"), which is unique per clause and whose choice is cross-referenced on the verb, from nominative-marking, which may apply to more than one argument including preverbal non-pivot topics.

Such examples lead us to suspect that the availability of non-pivot topics, especially with non-pivot agents, may be quite widespread across Philippine-type voice system languages. At the same time, this possibility appears not to be universal across these languages: an anonymous reviewer notes that non-pivot agent topicalization is not tolerated in Malagasy, another Philippine-type language, as recently documented in Xu 2019. Within the framework for extraction restrictions developed and defended here, we hypothesize that such variation reflects different featural specifications on the probes involved: if topicalization and clefting both involve the same probe specification, we would expect both constructions to exhibit the same extraction restriction.

The possibility of this cross-linguistic variation parallels the variation observed between different Ā-dependencies within Bikol, where clefting involves [PROBE:FOC+D] limited to the closest DP (Branan and Erlewine, to appearb) and topicalization involves [PROBE:TOP]. Although Ā-dependencies are often described as a natural class (see e.g. Chomsky, 1977), subsequent work has also shown that there are important distinctions in this space as well (see e.g. Cinque, 1990; Lasnik and Stowell, 1991; Postal, 1994). The featural specifications of the probes involved is one important way in which different Ā-dependencies are distinguished.

Examples of the form in (80) in Tagalog and the many other languages cited above have largely been ignored in previous discussions of Austronesian syntax, but we believe that they are important data points that show that the characterization of all Ā-dependencies in these languages as strictly pivot-only is overly simplistic. The careful investigation of such extraction restrictions — both between different languages as well as between different Ā-constructions in individual languages — will contribute to our broader understanding of the shape of possible variation in Ā-probing.

References

Aldridge, Edith. 2004. Ergativity and word order in Austronesian languages. Doctoral Dissertation, Cornell University.

Aldridge, Edith. 2008. Generative approaches to ergativity. *Language and Linguistics Compass* 10:966–995.

Aldridge, Edith. 2012. Antipassive and ergativity in Tagalog. *Lingua* 122:192–203.

Aldridge, Edith. 2017. φ-feature competition: A unified approach to the Austronesian extraction restriction. In *Proceedings of CLS 52*, ed. Jessica Kantarovich, Tran Truong, and Orest Xherija, 1–20.

Aldridge, Edith Catherine. 2002. Nominalization and *wh*-movement in Seediq and Tagalog. *Language and Linguistics* 3.

Baker, Mark C. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.

Baker, Mark C. 2015. *Case: Its principles and its parameters*. Cambridge University Press.

Belletti, Adriana. 2012. Revisiting the CP of clefts. In *Discourse and grammar: from sentence types to lexical categories*, ed. Günther Grewendorf and Thomas Ede Zimmermann, 91–114. De Gruyter Mouton.

Belletti, Adriana, and Chris Collins, ed. 2021. *Smuggling in syntax*. Oxford University Press.

Benton, Richard A. 1971. *Pangasinan reference grammar*. University of Hawaii Press.

Blake, A. L. 2020. Left-dislocation in Western Subanon. *Philippine Journal of Linguistics* 51:30–54.

Bodily, Jarren. 2003. Locative and benefactive voice constructions: a look at preposition incorporation. In *LACUS Forum 30*, 259–269.

Branan, Kenyon. to appear. Domains of ordering. *Journal of East Asian Linguistics* .

Branan, Kenyon, and Michael Yoshitaka Erlewine. to appeara. Locality and (minimal) search. In *Cambridge handbook of minimalism*, ed. Kleanthes K. Grohmann and Evelina Leivada. Cambridge University Press.

Branan, Kenyon, and Michael Yoshitaka Erlewine. to appearb. Ā-probing for the closest DP. *Linguistic Inquiry* .

Büring, Daniel. 2005. *Binding theory*. Cambridge University Press.

Ceña, Resty Mendoza, and Ricardo Ma Duran Nolasco. 2011. *Gramatikang Filipino: Balangkasan*. The University of the Philippines Press.

Chen, Tingchun. 2018. Multiple case assignment: An Amis case study. Doctoral Dissertation, Massachusetts Institute of Technology.

Chen, Victoria. 2017. A reexamination of the Philippine-type voice system and its implications for Austronesian primary-level subgrouping. Doctoral Dissertation, University of Hawai'i.

Chomsky, Noam. 1977. On *wh*-movement. In *Formal syntax*, ed. Peter Culicover, Thomas Wasow, and Adrian Akmajian, 71–132. New York: Academic Press.

Chomsky, Noam. 2000. Minimalist inquiries: the framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 89–156. MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 1–52. MIT Press.

Chow, Keng Ji. 2022. A novel algorithm for minimal search. *Snippets* 42:3–5.

Cinque, Guglielmo. 1990. *Types of A'-dependencies*. MIT Press.

Collins, Chris. 2005. A smuggling approach to the passive in English. *Syntax* 8:81–120.

Collins, James N. 2019. Definiteness determined by syntax: A case study in Tagalog. *Natural Language & Linguistic Theory* 37:1367–1420.

Constant, Noah. 2014. Contrastive topic: Meanings and realizations. Doctoral Dissertation, University of Massachusetts Amherst.

De Guzman, Videa P. 1988. Ergative analysis for Philippine languages: an analysis. In *Studies in Austronesian linguistics*, ed. Richard McGinn, 323–345. Ohio University Press.

De Guzman, Videa P. 1995. Experiencer verbs in Tagalog. In *Grammatical relations: Theoretical approaches to empirical questions*, ed. Clifford S. Burgess, Katarzyna Dziwirek, and Donna B. Gerdts, 45–62. CSLI Publications.

Deal, Amy Rose. 2015. Interaction and satisfaction in ϕ-agreement. In *Proceedings of NELS 45*, ed. Thuy Bui and Deniz Özyıldız, 1–14.

Deal, Amy Rose. 2016. Syntactic ergativity: Analysis and identification. *Annual Review of Linguistics* 2:1–21.

Deal, Amy Rose. 2017. Syntactic ergativity as case discrimination. In *Proceedings of WCCFL 34*, ed. Aaron Kaplan, Abby Kaplan, Miranda K. McCarvel, and Edward J. Rubin, 141–150.

Deal, Amy Rose. to appear. Interaction, satisfaction, and the PCC. *Linguistic Inquiry* .

Den Dikken, Marcel. 2018. *Dependency and directionality*. Cambridge University Press.

Donohue, Cathryn, and Mark Donohue. 2010. The case of posessors and 'subjects'. In *Austronesian and theoretical linguistics*, ed. Raphael Mercado, Eric Potsdam, and Lisa deMena Travis, 103–116. John Benjamins Publishing Company.

Douglas, Jamie. 2018. Māori subject extraction. *Glossa* 3:1–34.

Elomari, Abdessalam. 1998. The unoptionality of resumptive pronouns: The case of Moroccan Arabic. In *Perspectives on Arabic linguistics XI*, ed. Elabbas Benmamoun, Mushira Eid, and Niloofar Haeri, 47–64. John Benjamins.

Erlewine, Michael Yoshitaka. 2014. Subject marking on non-subjects in Squliq Atayal. Manuscript, to appear in *Proceedings of AFLA 20*.

Erlewine, Michael Yoshitaka. 2018. Extraction and licensing in Toba Batak. *Language* 94:662–697.

Erlewine, Michael Yoshitaka, and Theodore Levin. 2021. Philippine clitic pronouns and the lower phase edge. *Linguistic Inquiry* 52:408–425.

Erlewine, Michael Yoshitaka, Theodore Levin, and Coppe van Urk. 2015. What makes a voice system? On the relationship between voice marking and case. In *AFLA 21: The Proceedings of the 21st Meeting*

of the Austronesian Formal Linguistics Association, ed. Amber Camp, Yuko Otsuka, Claire Stabile, and Nozomi Tanaka, 51–68. Asia-Pacific Linguistics.

Erlewine, Michael Yoshitaka, Theodore Levin, and Coppe van Urk. 2017. Ergativity and Austronesian-type voice systems. In *Oxford Handbook of Ergativity*, ed. Jessica Coon, Diane Massam, and Lisa deMena Travis, 373–396. Oxford University Press.

Erlewine, Michael Yoshitaka, Theodore Levin, and Coppe van Urk. 2020. The typology of nominal licensing in Austronesian voice system languages. In *Proceedings of AFLA 26*, ed. Ileana Paul, 71–87.

Ferreirinho, Naomi. 1993. *Selected topics in the grammar of Limos Kalinga, the Philippines*. Pacific Linguistics.

Finer, Daniel L. 1997. Contrasting Ā-dependencies in Selayarese. *Natural Language & Linguistic Theory* 15:677–728.

Foley, William A. 2008. The place of Philippine languages in a typology of voice systems. In *Voice and grammatical relations in Austronesian languages*, ed. Peter K. Austin and Simon Musgrave, 22–44. CSLI Publications.

Fowlie, Meaghan. 2013. Multiple multiple spellout. In *Challenges to linearization*, ed. Theresa Biberauer and Ian Roberts, 129–169. Mouton de Gruyter.

Gerdts, Donna B. 1988. Antipassives and causatives in Ilokano: evidence for an ergative analysis. In *Studies in Austronesian linguistics*, ed. Richard McGinn, 295–321. Ohio University Press.

Guilfoyle, Eithne, Henrietta Hung, and Lisa Travis. 1992. Spec of IP and Spec of VP: Two subjects in Austronesian languages. *Natural Language & Linguistic Theory* 10:375–414.

Halpert, Claire. 2012. Argument licensing and agreement in Zulu. Doctoral Dissertation, Massachusetts Institute of Technology.

Halpert, Claire. 2016. Raising parameters. In *Proceedings of WCCFL 33*, 186–195.

Halpert, Claire. 2019. Raising, unphased. *Natural Language & Linguistic Theory* 37:123–165.

Hsieh, Henrison. 2019. Distinguishing nouns and verbs: A Tagalog case study. *Natural Language & Linguistic Theory* 37:523–569.

Hsieh, Henrison. 2020. Beyond nominative: A broader view of A'-dependencies in Tagalog. Doctoral Dissertation, McGill University.

Hsieh, Henrison. 2021. Locality in exceptional Tagalog A'-extraction. Manuscript, National University of Singapore.

Iatridou, Sabine. 1995. Clitics and island effects. In *Penn working papers in linguistics*, volume 2.

Imanishi, Yusuke. 2019. The clause-mate condition on resumption: Evidence from Kaqchikel. *Studia Linguistica* 73:398–441.

Kaufman, Daniel. 2010. The morphosyntax of Tagalog clitics: A typologically driven approach. Doctoral Dissertation, Cornell University.

Kaufman, Daniel. 2017. Lexical category and alignment in Austronesian. In *The Oxford handbook of Ergativity*, ed. Jessica Coon, Diane Massam, and Lisa Travis, 589–628. Oxford University Press.

Keenan, Edward. 1976. Remarkable subjects in Malagasy. In *Subject and topic*, ed. Charles N. Li and Sandra A. Thompson. Academic Press.

Keenan, Edward L., and Bernard Comrie. 1977. Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry* 8:63–99.

Kroeger, Paul R. 1990. Case incorporation in Philippine languages. Manuscript, Stanford.

Kroeger, Paul R. 1991. Phrase structure and grammatical relations in Tagalog. Doctoral Dissertation, Stanford University.

Kroeger, Paul R. 1993. Another look at subjecthood in Tagalog. *Philippine Journal of Linguistics* 24:1–16.

Lasnik, Howard, and Tim Stowell. 1991. Weakest crossover. *Linguistic Inquiry* 22:687–720.

Latrouite, Anja. 2011a. Differential object marking in Tagalog. In *Proceedings of AFLA 18*, ed. Lauren Eby Clemens, Gregory Scontras, and Maria Polinsky, 94–109.

Latrouite, Anja. 2011b. Voice and case in Tagalog: The coding of prominence and orientation. Doctoral Dissertation, Heinrich-Heine Universität.

Levin, Theodore. 2015. Licensing without case. Doctoral Dissertation, Massachusetts Institute of Technology.

Marantz, Alec. 1984. *On the nature of grammatical relations*. MIT Press.

Marantz, Alec. 1991. Case and licensing. In *Proceedings of ESCOL 8*, ed. Germán F. Westphal, Benjamin Ao, and Hee-Rahk Chae, 234–253.

McCloskey, James. 1979. *Transformational syntax and model theoretic semantics: A case study in Modern Irish*. Dordrecht: Reidel.

McCloskey, James. 2002. Resumption, successive cyclicity, and the locality of operations. In *Derivation and explanation in the minimalist program*, ed. Samuel David Epstein and T. Daniel Seely, 184–226. Blackwell.

McFarland, Curtis Daniel. 1974. The dialects of the Bikol area. Doctoral Dissertation, Yale University.

Miller, Barry W. 1988. Non-configurationality in Tagalog. Doctoral Dissertation, University of Michigan.

Mintz, Malcolm W. 2019. *Bikol dictionary*. University of Hawai'i Press.

Mintz, Malcolm Warren. 1973. Case and semantic affixes of Bikol verbs. Doctoral Dissertation, University of Hawai'i.

Mirikitani, Leatrice T. 1972. *Kapampangan syntax*. University of Hawaii Press.

Mithun, Marianne. 1994. The implications of ergativity for a Philippine voice system. In *Voice: Form and function*, ed. Barbara A. Fox and Paul J. Hopper, number 27 in Typological Studies in Language, 247–277. John Benjamins.

Mithun, Marianne. 2019. Grammatical relations in Hiligaynon. In *Argument selectors: A new perspective on grammatical relations*, ed. Alena Witzlack-Makarevich and Balthazar Bickel, 131–184. John Benjamins.

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

Payne, Thomas E. 1982. Role and reference related subject properties and ergativity in Yup'ik Eskimo and Tagalog. *Studies in Language* 6:75–106.

Pizarro-Guevara, Jed Sam. 2020. When human universal meets language specific. Doctoral Dissertation, University of California Santa Cruz.

Polinsky, Maria. 2017. Syntactic ergativity. In *The Blackwell companion to syntax, 2nd edition*, ed. Martin Everaert and Henk van Riemsdijk. Blackwell.

Postal, Paul M. 1994. Contrasting extraction types. *Journal of Linguistics* 30:159–186.

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

Rackowski, Andrea. 2002. The structure of Tagalog: Specificity, voice, and the distribution of arguments. Doctoral Dissertation, Massachusetts Institute of Technology.

Rackowski, Andrea, and Norvin Richards. 2005. Phase edge and extraction: a Tagalog case study. *Linguistic Inquiry* 36:565–599.

Reid, Lawrence A. 1978. Problems in the reconstruction of Proto-Philippine construction markers. In *Proceedings of ICAL 2*, volume 1, 33–66. Pacific Linguistics.

Reid, Lawrence A., and Hsiu-chuan Liao. 2004. A brief syntactic typology of Philippine languages. *Language and Linguistics* 5:433–490.

Richards, Norvin. 2005. Person-case effects in Tagalog and the nature of long-distance extraction. In *Proceedings of AFLA 12*, ed. Jeffrey Heinz and Dimitrios Ntelitheos, 383–394.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In *Elements of grammar*, ed. Liliane Haegeman, 281–337. Kluwer Academic Publishers.

Ross, John Robert. 1967. Constraints on variables in syntax. Doctoral Dissertation, Massachusetts Institute of Technology.

Ross, Malcolm. 2009. Proto Austronesian verbal morphology: A reappraisal. In *Austronesian historical linguistics and culture history: A festschrift for Robert Blust*, ed. Alexander Adelaar and Andrew Pawley, 295–326. Canberra: Pacific Linguistics.

Rowsell, Lorna V. 1983. An ergative analysis of Kapampangan. Master's thesis, University of Calgary.

Sabbagh, Joseph. 2016. Specificity and objecthood in Tagalog. *Journal of Linguistics* 52:639–688.

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

Schachter, Paul. 1996. The subject in Tagalog: still none of the above. In *UCLA Occasional Papers in Linguistics*, volume 15, 1–61. Los Angeles: Department of Linguistics, UCLA.

Schachter, Paul, and Fe T. Otanes. 1972. *Tagalog reference grammar*. University of California Press.

Schütze, Carson T. 2001. On the nature of default case. *Syntax* 2:205–238.

Shibatani, Masayoshi. 1988. Voice in Philippine languages. In *Passive and voice*, ed. Masayoshi Shibatani, 85–142. John Benjamins.

Shibatani, Masayoshi. 1991. Grammaticization of topic into subject. In *Approaches to grammatiaclization*, ed. Elizabeth Closs Traugott and Bernd Heine, volume 2, 93–133. John Benjamins.

Shlonsky, Ur. 1987. Focus constructions in Berber. In *Studies in Berber syntax*, ed. Mohamed Guerssel and Kenneth Hale, 1–20. MIT Working Papers in Linguistics.

Sportiche, Dominique. 2017. Relative clauses: Promotion only, in steps. Manuscript, UCLA.

Sportiche, Dominique. 2018. Resumed phrases (are always moved, even with in-island resumption). In *Romance languages and linguistic theory 14*, ed. Francisco Ordóñez and Lori Repetti. John Benjamins.

Tsukida, Naomi. 2018. Preposed NPs in Seediq. In *Perspectives on information structure in Austronesian languages*, ed. Sonja Riesberg, Asako Shiohara, and Atsuko Utsumi, 313–344. Language Science Press.

van Urk, Coppe. 2015. A uniform syntax for phrasal movement: A Dinka Bor case study. Doctoral Dissertation, Massachusetts Institute of Technology.

van Urk, Coppe. 2017. Phase impenetrability and resumption in Dinka. Presented at GLOW 40.

van Urk, Coppe, and Norvin Richards. 2015. Two components of long-distance extraction: Successive cyclicity in Dinka. *Linguistic Inquiry* 46:113–155.

Xu, Kang. 2019. On the topic of Malagasy *dia* topicalization construction. Master's thesis, University of Western Ontario.