Ā-probing for the closest DP

Kenyon Branan & Michael Yoshitaka Erlewine

Draft, August 2021

Abstract We consider the typology of attested Ā-extraction asymmetries between core argument DPs and argue that an Ā-probe can be required to specifically target the closest accessible DP. Such an Ā-probe specification is part of the influential Aldridge 2004, 2008 analysis of syntactically ergative extraction restrictions, but has not been widely adopted outside of work on ergative languages. We argue that restricted probing of this form underlies subject-only extraction behaviors in a number of non-ergative languages, including some of those in Keenan and Comrie's (1977) typology of relativization asymmetries. We describe the behaviors of such probes in detail and relate them to other probe-goal behaviors in recent work on composite A/Ā probes.

Acknowledgements For discussion of specific facts reported in the paper, we thank Henrison Hsieh (Tagalog) and Jaklin Kornfilt (Turkish). For additional discussion related to other languages in the Keenan & Comrie survey which are unfortunately not discussed here due to space constraints, we thank Dan Brodkin (Minangkabau), Catriona Hyslop Malau (Northeast Aoban/Ambae), and Alex Smith and Carly Sommerlot (Iban). We also thank Amy Rose Deal and Asia Pietraszko for related discussion, and Justin Adhiyatma for preparing our Rejang corpus materials. This research is supported by the Singapore Ministry of Education under grant MOE2017-T2-2-094 "Subjecthood in Southeast Asia: Description and Theory" to the second author.

Contents

1 Introduction

A central concern of syntactic theory is how non-local dependencies are formed, and how they are constrained. Since Chomsky 2000, 2001, much of this work has been fruitfully discussed in terms of *probes* and their specifications. Probes initiate a search for a *goal* that matches a particular feature specification, to Agree with or to Move. This paper contributes to the question of the possible feature specifications and behaviors of probes that trigger Ā-movement.

For example, we may describe *wh*-movement in a language like English as involving C probing for the closest [WH] constituent (see e.g. Rizzi, 1990). This allows for *wh*-movement of the embedded subject in (1a) or the embedded object in (1b); in either case, the moved goal is the closest constituent with a [WH] feature. Intervening non-*wh* constituents are ignored. When there are multiple potential goals accessible to the probe, the structurally closest goal is chosen, as reflected in the contrast in (1c).

(1) **Ā-probing for the closest [WH] goal:**

a. C[PROBE:WH] you expect *who* to eat the sandwich ⇒
Who do you expect ___ to eat the sandwich?

b. C[PROBE:WH] you expect Sara to eat *what* ⇒
What do you expect Sara to eat ___?

c. C[PROBE:WH] you expect *who* to eat *what* ⇒
Who do you expect ___ to eat *what*?
**What* do you expect *who* to eat ___?

Concretely, we adopt the definition for "closest" in (2), which is equivalent to the locality condition on Agree stated in Chomsky 2000: 122:

(2) **Closest:**
A potential goal G for probe P is closest if no other potential goal for P c-commands G.

There are, however, languages with much stricter restrictions on Ā-extraction, such that only particular types of arguments can be Ā-extracted. Aldridge 2004, 2008 develops one influential

approach to the analysis of so-called syntactic ergativity, narrowly defined as a ban on the Ā-extraction of transitive subjects. (See Deal 2015, 2016 and Polinsky 2017 for recent overviews and discussion.) One component of Aldridge's analysis is a claim that Ā-probing can be restricted as in (3):

(3) **Ā-probing for the closest DP:**

An Ā-probe can be specified to target the closest accessible DP.

Combined with a commonly adopted approach to the clause structure of a sub-type of ergative languages, a probe of this type will necessarily target absolutive arguments. Ā-probing of the form in (3) has been adopted for the analysis of Philippine-type Austronesian languages, for which the analysis was developed, as well as for similar extraction restrictions in Mayan languages in more recent work (Levin, 2018; Coon, Baier, and Levin, 2021). However, probing of this form has not been commonly invoked in the analysis of non-ergative languages, leading to an impression by some that probing of this form is a special property of ergative languages — just as it has been claimed that syntactically ergative extraction restrictions exist but syntactically accusative ones do not.

In this paper, we argue that the Ā-probing for the closest DP (3) is indeed attested in the grammars of non-ergative languages, manifest in extraction constructions with apparent subject-only restrictions. In particular, Ā-probing for the closest DP makes accurate predictions for apparent exceptions to subject-only extraction restrictions: for example, even in a language where subjects are frequently the highest DP in a clause, if the language has a strategy for raising a non-subject DP to a higher position, it may feed the restricted extraction. We conclude that Ā-probes indeed can be specified to necessarily target the closest DP, as proposed by Aldridge, and that such Ā-probes are not limited to ergative languages.

After reviewing the motivation for this conjecture as part of the analysis of syntactic ergativity in section 2, we formalize this mode of probing and present two novel arguments for it from relativization in Turkish and Rejang in section 3. In section 4, we then review and highlight relevant results from the Keenan and Comrie 1977 *et seq* typology of relativization and related subsequent work, which serves to motivate and contextualize the current work. Finally, we argue that this restriction on Ā-probing to the closest DP must be a specification on individual probes, rather than a language-level or construction-level parameter in section 5 and conclude in section 6.

2 Syntactic ergativity via Ā-probing for the closest DP

We begin by reviewing Aldridge's (2004, 2008) analysis for syntactic ergativity in so-called "Philippine-type" Austronesian languages, such as Tagalog. Philippine-type languages are verb-initial with case marking patterns which can be analyzed as exhibiting ergative-absolutive alignment.[1] Among core arguments of the verb, these languages allow only for Ā-extraction of the absolutive DP:

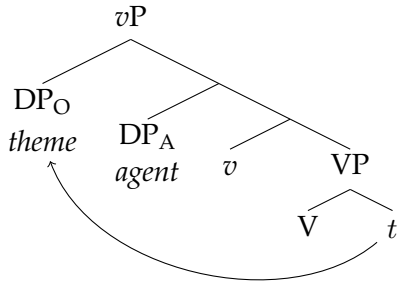
(4) **Absolutive-only extraction restriction in Tagalog:** (Henrison Hsieh p.c.)

a. tela=ng [b<in>ili ng bata]
cloth=LK <PRF>buy ERG child
'cloth that the child bought'

b. *bata=ng [b<in>ili ang tela]
child=LK <PRF>buy ABS cloth
'child who bought cloth'

Aldridge's theory for this extraction restriction is one specific instantiation of what Deal 2016 refers to as the "standard theory of syntactic ergativity," narrowly referring to the absolutive-only extraction restriction. The shared intuition of these proposals is that transitive objects (O) canonically occupy a structural position above that of transitive subjects (A).[2] For Aldridge, in a transitive clause with two core arguments, an agent and theme, the theme will move to an outer specifier of *v*P.[3] The agent is base-generated as the inner specifier of DP. See (5). The verb is ultimately pronounced higher, preceding its arguments.

(5) **Monotransitive *v*P as in Aldridge's account:**



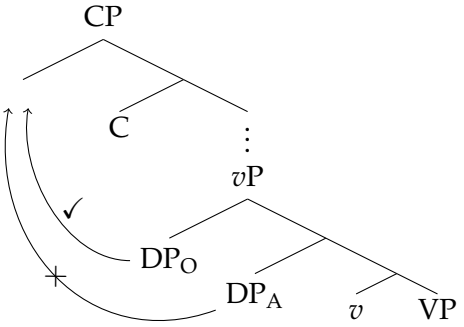
[1] "Philippine-type" refers to a set of languages with certain shared grammatical characteristics; see Himmelmann 2002, Ross 2002, Blust 2010. Another major subgroup of Austronesian languages is the "Indonesian-type," discussed below. The description of Philippine-type Austronesian languages as ergative has however been controversial. See for example Erlewine et al. 2017 and Chen 2017.

[2] Other examples of the "standard theory" include Campana 1992, Murasugi 1992, Ordóñez 1995, Bittner and Hale 1996, Coon et al. 2014, Assmann et al. 2015, Levin 2018, and Coon et al. 2021.

Aldridge also discusses clauses with applicatives, where the DP moved to the outer specifier of *v*P is a goal, instrument, location, or possessor (see Nie, 2019) instead of a theme.

We now turn to the question of Ā-extraction. Ā-extraction of DPs in Tagalog is limited to the extraction of absolutive arguments, e.g. transitive objects (O) and intransitive subjects (S); transitive subjects (A) cannot be Ā-extracted. Aldridge proposes that Ā-probing by C necessarily targets the closest DP.[4] Following the proposed structure for transitive clauses, Ā-extraction from a transitive clause will thus necessarily target the outer specifier of *v*P, which may be a transitive object (O) or an applicativized argument (see above). See (6). There is no way to target a transitive subject (A) for Ā-movement.[5]

(6) **Ā-probing for the closest DP from (5):**



In an intransitive clause, the sole DP argument (S) is closest to the probe and thus can be Ā-extracted as well.[6] This derives the syntactically ergative Ā-extraction restriction.

---

[3] This analysis dovetails with the widely-adopted inherent case theory for ergative case (e.g. Woolford, 1997, 2006; Legate, 2002, 2008; Aldridge, 2004, 2008; see also Sheehan 2017 for a recent overview). Because the agent receives ergative case in its thematic position, there is no need for the agent to syntactically associate with a higher functional head such as T as in many proposals for structural nominative case.

[4] Specifically, Aldridge (2004: 338) writes: "C has an EPP feature, which attracts a DP. In a transitive clause, the closest DP will be the internal argument absolutive, residing in the outer specifier of v. The external argument will not be attracted, because doing so would violate Attract Closest." Aldridge (2012: 197 fn 9) later clarifies that "closest" must be defined so that two specifiers of the same phrase do not count as equidistant for higher probes, *pace* Chomsky's (2000: 122, 130; 2001: 27) Equidistance principle. Our definition of "closest" in (2) satisfies this desideratum.

[5] Syntactically ergative languages generally have a strategy for Ā-extracting notional transitive subjects (A). A common one is to antipassive the clause, so that the A subject becomes a formally intransitive S subject, and thus eligible for Ā-extraction. See Aldridge 2012 for discussion of this approach in Tagalog.

[6] Intransitive *v* for Aldridge does not move any argument to its specifier. Intransitives are either unergative, with the

Aldridge's analysis for the extraction asymmetry in Philippine-type languages thus relies on the conjecture stated in (3), repeated here:

(7) **Ā-probing for the closest DP:** =(3)

An Ā-probe can be specified to target the closest accessible DP.

This conjecture in (7) has been adopted as part of the analysis of syntactically ergative extraction asymmetries in other languages as well, including recently in Levin 2018 and Coon, Baier, and Levin 2021 for a number of Mayan languages. We also refer the reader to Erlewine and Lim 2019 for an investigation of extraction asymmetries in Bikol, a sister language to Tagalog, which strengthens the empirical case for extraction asymmetries in Philippine-type Austronesian languages to be based on (7), over and above Aldridge's original argumentation.

We should however step back and note that Ā-probing of this form — if it exists — is perhaps unusual and conceptually surprising.[7] Such a probe would lead to an Ā-extraction process that has the locality profile of A-movement, rather than the familiar long-distance and 'relativized' character of Ā-movement (as in e.g. Chomsky, 1977; Rizzi, 1990; see (1) above). It also raises questions for whether and how non-DPs can be Ā-extracted in such a language, which we return to at the end of this paper. We also note that alternative accounts for the extraction behavior of such Philippine-type languages exist, which do not require Ā-probing for the closest DP (7).[8] In the pursuit of a maximally restrictive theory of grammar, then, it is tempting to reject the possibility of probing of the form in (7), or perhaps to somehow limit its availability to ergative languages.

The goal of this paper is to show that there in fact *is* substantial motivation for the existence of Ā-probing that is restricted to the closest DP (7), in non-ergative languages. This conclusion in turn lends support for the plausibility of Aldridge's proposal for absolutive-only extraction restrictions in ergative languages as well.

---

agent DP being the sole specifier of *v*P, or unaccusative, where the sole DP argument is lower, but with *v*P being a "weak phase" in Chomsky's (2001) terms and thus permeable for probing from above.

[7] There is rather little explicit discussion of this notable aspect of Aldridge's proposal. We are aware of such discussion only in Aldridge 2008: 990, 992 note 6, Deal 2015: 698–699, and Polinsky 2017: 18–20.

[8] For example, see the "case agreement" approaches of Pearson 2001, 2005 for Malagasy and Rackowski 2002 and Rackowski and Richards 2005 for Tagalog, which take the apparent absolutive-only extraction restriction to be epiphenomenal. Another alternative approach discusses apparent "extraction" restrictions in these languages without appealing to

3 Subject-only extraction restrictions from Ā-probing for the closest DP

Given the potentially unusual nature of the idea that an Ā-probe would be limited to attracting the closest DP goal (7), as discussed above, we seek independent motivation for this possibility in grammar from beyond its original application to syntactic ergativity. In this section we present two case studies of relativization in two non-ergative languages, Turkish and Rejang, which we claim are best analyzed as involving an Ā-probe that can only attract the structurally closest DP. This approach will allow us to account not only for the basic subject-only restriction on these Ā-processes, but also for its apparent exceptions.

Let us begin by detailing, in schematic terms, the behavior that we expect from Ā-probing for the closest DP (7).[9] Here we describe such probes as a particular version of a probe that seeks a goal that bears both an Ā-feature (e.g. WH, REL) and an A-feature (D), which we notate [PROBE:Ā+D] in the general case. This probe is additionally required to match with a goal that is the closest DP to the probe, with "closest" defined as in the discussion of Agree in Chomsky 2000: 122, repeated here from (2) above:

(8) **Closest:**

A potential goal G for probe P is closest if no other potential goal for P c-commands G.

With [PROBE:Ā+D] restricted to the closest DP, we expect extraction of DP_2 in (9) to be impossible: DP_1 c-commands DP_2 and therefore counts as structurally closer to the probe.

(9) * [PROBE:Ā+D] ... [DP_1 ... [$DP_{2[\bar{A}]}$ ...

extraction at all, instead analyzing different "voice" forms as different participant nominalizations; see Keenan 2008 for Malagasy and Kaufman 2009 for Tagalog. (But see also Hsieh 2019 for a forceful response to Kaufman.) Such nominalization analyses may also extend to subject-only participial relatives in European languages mentioned by Keenan and Comrie (1977: 70); see fn. 24 below.

For syntactic ergativity in other language families, too, there are accounts which do not involve Ā-probing for the closest DP (7). See Deal 2016, 2017 and Polinsky 2017 for two recent approaches.

Here we will not review or evaluate the arguments for analyzing syntactically ergative extraction restrictions in (potentially) morphologically ergative languages as involving Ā-probing for the closest DP and instead concentrate on the applicability of Ā-probing for the closest DP in non-ergative languages.

[9] See Branan and Erlewine forthcoming for discussion of how probing of this form can be implemented in procedural terms.

This logic makes two predictions. First, if a language has an independent mechanism for bringing DP_2 above DP_1, we expect the probe to then be able to interact with DP_2:

(10) ✓[PROBE:Ā+D] ... [$DP_{2[\bar{A}]}$ ... [DP_1 ... [<$DP_{2[\bar{A}]}$> ...

Second, if there is no c-command relationship between two DPs, we might expect either to be extractable, as neither is closer to the probe than the other. (See again the definition of "closest" in (8).) One such configuration would arise when DP_2 is at the edge of DP_1:

(11) ✓[PROBE:Ā+D] ... [$_{DP1}$ $DP_{2[\bar{A}]}$ ...] ...

Two constituents in this configuration both counting as "closest" for a higher probe is also discussed and invoked in Pesetsky and Torrego 2001: 363.

We now present evidence from relativization in Turkish and Rejang that motivate the existence of Ā-probing for the closest DP. In particular, we will see that the configurations in (10) and (11) indeed allow for Ā-extraction of DP_2 in constructions that disallow extraction in the configuration in (9). Neither language is morphologically ergative. We also situate our proposal in relation to recent work (e.g. Van Urk and Richards, 2015; Bossi and Diercks, 2019; Colley and Privoznov, 2020; Scott, 2021) which argues for probes of the form [PROBE:Ā+D] yet allow said probes to skip closer partial matches.

3.1 Turkish

Turkish has two forms for relative clauses, shown below in (12), traditionally described as a subject/non-subject distinction (Underhill, 1972; Hankamer and Knecht, 1976; a.o.). We follow Cagri 2005, 2009 in glossing these forms 'subject relative' (SR) and 'non-subject relative' (NSR) here. Here we concentrate on the behavior of relativization with the SR suffix, which receives a simple analysis as the exponent of a head which has an Ā-probe which is specified to probe for the closest DP. We will also later discuss its relation to NSR forms.

(12) **Two relative clause forms in Turkish:**

a. [___ kız-ı sok-**an**] arı
girl-ACC sting-SR bee
'the bee that stung the girl'
(Cagri, 2005: 24 ex. 15a)

b. [arı-nın ___ sok-**tuğ**-u] kız
bee-GEN sting-NSR-3SG girl
'the girl that the bee stung'
(Jaklin Kornfilt, p.c.)

It is well known that, under certain limited circumstances, non-subjects may be extracted using the SR form (see e.g. Underhill 1972, Hankamer and Knecht 1976, Kornfilt 1984, 1997, 2000, and subsequent work). For instance, object relativization with the SR form is possible in examples such as (13a), where the subject is indefinite and in a low, immediately preverbal position.[10] The source for this relative clause in (13a) is therefore (13b), where the accusative object is the highest DP in the clause. In contrast, object relativization across a high, non-indefinite subject must use the NSR form, as in (12b) above.

(13) **Low indefinite subject makes object highest:**

(Temürcü, 2001: 147 ex. 199a, 146 ex. 197a)

a. Object relative with SR form:
[___ arı sok-an] adam
bee sting-SR man
'the man stung by a bee'

b. Source structure:
Adam-ı arı sok-tu.
man-ACC bee sting-PAST
'A bee stung the man.'

Temporal relatives may also be formed with SR if the subject is a low, indefinite subject, as in (14) below. This too is explained by our analysis for the SR probe as temporal adjunct DPs naturally occupy a position above low indefinite subjects.

[10] Evidence for the low structural position of indefinite immediately preverbal subjects comes from their position with respect to low adverbs. Taylan (1984) shows that morphologically simple manner adverbs must be immediately preverbal, except in the case where a caseless indefinite object follows it. Such adverbs have been used as a diagnostic for the edge of the predicate domain (Kornfilt, 1984; Kelepir, 2001; Öztürk, 2009). Low indefinite subjects also follow these adverbs:

(i) Ali-yi fena arı sok-tu.
Ali-ACC badly bee sting-PAST
'A bee stung Ali badly.' (Sağ-Parvardeh, 2019: 92 fn 20)

(14) **Temporal relatives with SR form:** (Cagri, 2005: 180 ex. 62a,b)

a. [___ bomba patlay-an] gün
bomb explode-SR day
'the day a bomb exploded'

b. [___ kar yağ-an] gün-ler
snow rain-SR day-PL
'the days it snowed'

Such examples show that the apparent subject orientation of the SR form is really a requirement to extract the closest nominal, rather than to extract the subject per se. We propose that SR reflects the use of [PROBE:REL+D] restricted to extract only the closest nominal in its domain. Movement of the object (DP_2 in (15)) across a low, indefinite subject (DP_1) allows for the SR probe to target the object:

(15) ✓[PROBE:REL+D] ... $DP_{2[REL]}$ ... [DP_1 ... <DP_2>

Possessor relativization presents a further class of apparent exceptions to the subject orientation of SR relatives. Example (16) shows that the possessor of a subject is extracted using the SR form. Note that subjects in NSR clauses as in (16b) receive genitive case, as seen in the grammatical (12b) above, whereas subjects of SR clauses as in (16a) retain their unmarked, nominative case.

(16) **Possessor-of-subject relative with SR form:** (Underhill, 1972: 88 ex. 4, 89 ex. 6)

a. [[___ oğl-u] mekteb-e gid-en] adam
son-POSS.3SG school-DAT go-SR man
'the man whose son goes to school'

b. *[[___ oğl-u-un] mekteb-e git-tiğ-i] adam
son-POSS.3SG-GEN school-DAT go-NSR-3SG man

This possibility of possessor relativization is also explained by our analysis of SR involving Ā-probing for the closest DP and our definition of "closest" in (8). Both the subject (DP_1 in (17)) as well as its possessor (DP_2) count as "closest" for the probe, as the subject DP_1 does not c-command its possessor DP_2.

(17) ✓[PROBE:REL+D] ... [$_{DP1}$ $DP_{2[REL]}$ ...] ...

Also as predicted by our approach, SR relativization can also target the possessor of an object if the subject is in the low, immediately preverbal position. See example (18). The derivation for (18) simply combines the object movement above the indefinite subject which stays low, illustrated in (15), with the SR probe targeting the possessor of the highest DP, which counts as "closest," as in (17) above. This derivation is illustrated in (19).[11]

(18) **Possessor-of-object relative with low subject, with SR form:** (*ibid.*: 133 ex. 36)

[[___ kız-ın-ı] arı {sok-an / *sok-tuğ-u}] adam

girl-POSS.3SG-ACC bee sting-SR / *sting-NSR-3SG man

'the man whose daughter a bee stung'

(19) ✓[PROBE:REL+D] ... [$_{DP2}$ $DP_{3[REL]}$ ...] ... [DP_1 ... <DP_2>

In contrast, if the subject is higher than the object — whether a bare definite or a *birkaç* 'some' indefinite — relativization of the object's possessor takes the NSR form:

(20) **Possessor-of-object relative with high subject, with NSR form:**

a. [(birkaç) arı-nın [___ bacağ-ın-ı] sok-tuğ-u] kız

some bee-GEN leg-POSS.3SG-ACC sting-NSR-3SG girl

'the girl whose leg the/some bee stung' (Cagri 2005: 99 ex. 29a; 32 ex. 25a)

b. *[(birkaç) arı [___ bacağ-ın-ı] sok-an] kız

some bee leg-POSS.3SG-ACC sting-SR girl

'the girl whose leg the/some bee stung' (*ibid.*: 99 ex. 29b and Jaklin Kornfilt, p.c.)

The behavior of relativization with the SR form shows the clear hallmarks of Ā-probing for the closest DP. SR relativization often targets subjects, which are generally highest in the clause, but can also target non-subjects when the subject is exceptionally low, as well as possessors of the highest DP which also count as closest. This suggests both that the regular subject orientation of SR is due

[11] This derivation constitutes an instance of "smuggling"; see Belletti and Collins 2021 for a recent overview. See also Nakamura 1996 and Branan 2018 for discussion of similar subextraction facts in Tagalog, compatible with Aldridge's account and the discussion of possessor subextraction here.

to subjects regularly being the highest DP in the clause, rather than SR specifically correlating with other subjecthood properties.[12]

Next, we address the structure of NSR relatives and their relation to SR relatives. The data above shows that SR and NSR forms are in complementary distribution, with SR appearing specifically where relativization could employ Ā-probing for the closest DP, and NSR appearing otherwise.[13] We propose that this reflects a preference for the use of SR relativization where possible, which in turn reflects a more general and familiar pressure for grammar to minimize probe-goal relations while maximizing featural matching.

Specifically, we propose in any relative clause, the complex probe we have described here, [PROBE:REL+D], on the relative complementizer probes first. If it finds a match, we extract this goal and the resulting complementizer is realized as SR morphology. If this joint probing does not find a match, the head is able to probe separately as [PROBE:REL] and [PROBE:D]; the former will attract a [REL] goal which was inaccessible using Ā-probing for the closest DP, and the latter will copy back subject φ-features. The resulting complementizer is then spelled out as the *-DIK* suffix followed by subject φ-agreement, which is more generally the default form for nominalized finite embedded clauses in the language. Here, the complex probe is privileged over two simple probes,

[12] For instance, SR morphology does not correlate with relativization of a nominal with a particular case form, either as the result of agreement in case features (Chung, 1982, 1994, 1998; Georgopoulos, 1985, 1991; Pearson, 2005; Rackowski and Richards, 2005) or by SR reflecting a case-discriminating probe (Deal, 2017). The SR probe's target is nominative in (12a), accusative in (13a), and genitive in (16, 18).

A reviewer asks whether the difference between the SR and NSR forms could be thought of as an alternation in complementizer morphology akin to the French *que*/*qui* alternation. (This connection is also made in Kornfilt 2008.) Here it is relevant to note that subextraction from a subject does not trigger the *qui* form of the complementizer in French:

(i) Combien est-ce que tu crois {que / *qui} [___ de gens] viendront?
how.many is-it that you think que / *qui of people will.come
'How many people do you think will come?' (Starke, 2001: 45)

This clearly contrasts with the behavior of SR morphology in Turkish, which is also employed for subextraction from the highest DP. Thus analyses for the *que*/*qui* alternation and similar, which generally do not involve distinct probes (see Pesetsky 2017 for an overview) cannot be directly imported for Turkish.

[13] Cagri (2005: 33 fn 36) notes that there is some dialectal variation in the strength of this effect. This could be analyzed as individual grammars differing in the strength of the preference we describe here below, or as reflecting the availability of distinct underlying structures for some speakers.

since a perfect match for the complex probes will check two features with only one probe-goal relationship.

Probe interactions very similar to this have been recently proposed by Van Urk and Richards (2015) for Dinka and Bossi and Diercks (2019) for Kipsigis. In each case, there is a probe which seeks a goal with a combination of A- and Ā-features. If there is an accessible goal which satisfies both these requirements simultaneously, the probe will target it, due to a principle that Van Urk and Richards (2015) call *Multitasking*. But if this joint probing fails, the head will instead probe separately for a goal with the A-feature and a goal with the Ā-feature.[14] There are just two differences between these interactions in Dinka and Kipsigis versus that in Turkish. First, the result of probing separately results in multiple movement in Dinka and Kipsigis, whereas in Turkish this results in movement of the [REL] goal and φ-agreement with the DP. Second, when probing jointly for a goal with both A- and Ā-features, in Turkish we must Ā-probe for the closest DP, whereas in Dinka and Kipsigis, closer partial matches can be skipped.[15] Abstracting away from these two differences, however, the mechanism that yields apparent "competition" between SR and NSR forms is equivalent to that proposed in these prior works.[16]

Finally, we note that the possibility for an unsatisfied complex probe such as [PROBE:Ā+D] to subsequently probe separately for Ā- and D-goals must be subject to lexical variation, rather than being a general property of complex probes. In particular, in the case studies we discuss below, as well as those in §5, we show that if Ā-probing for the closest DP is unable to find a suitable goal, the relevant Ā-construction is simply impossible in that configuration.

[14] Although not described in such terms, these interactions could also be productively described in terms of head or probe "splitting" (Martinović, 2015), or else their preferred "bundling" (Erlewine, 2018). See Erlewine 2018 sec. 4.4 (especially note 32) for discussion of such ideas in prior literature.

[15] See also Colley and Privoznov 2020 on Khanty and Scott 2021 on Ndengeleko for other case studies proposing probes which can target goals with both A- and Ā-features that are not the closest DP.

[16] This effect also echoes the *Fewest Steps* principle from Chomsky 1991, 1995 (see also Collins 2001 and references therein), which favors derivations with fewer steps. It may also be described as a violable version of *Maximize Matching Effects* from Chomsky 2001.

3.2 Rejang

Rejang is an Austronesian language spoken in southwest Sumatra. Ā-extraction of DPs in the language is limited to the highest accessible nominal, whether or not that is the subject.[17]

Rejang is canonically SVO, with transitive verbs having two verb forms, active and passive. Active themes and passive agents cannot be Ā-extracted — with one exception, discussed below. For example, transitive theme relativization requires first promoting the theme to subject position using a passive:

(21) **Rejang theme relativization requires passivization:** (McGinn, 1998: 362 ex. 5b, 6)

a. *tun [gi pelisi o m-akep ___ kelem] o
person C_{gi} police the ACT-catch last.night the
'the person that the police arrested last night'

b. tun [gi t<en>akep pelisi ___ kelem] o
person C_{gi} PASS-catch police last.night the
'the person that was arrested by the police last night.'

There is evidence that this apparent subject-only restriction on extraction in Rejang is in fact better described as a requirement that only the highest nominal may undergo extraction. Evidence for this comes from three sources: agent cliticization, long distance extraction, and extraction of possessors, which we discuss one by one.

McGinn (1998: 372ff) notes one exception to the ban on active theme extraction: Extraction of active themes is possible if the agent subject is a clitic pronoun, as in (22).[18]

(22) **Active theme *wh*-question across a pronominal subject:** (McGinn, 1989: 208 ex. 1b)

Jano [gi ko tokoa ___]?
what C_{gi} 2sg ACT-buy
'What did you buy?'

[17] Here we rely heavily on the work of Richard McGinn, who conducted extensive fieldwork on Rejang for over four decades, but unfortunately passed away in 2018. (See Blust 2018.) There are some empirical questions which we cannot answer based on these existing works; perhaps most importantly, there are no tests of island-sensitivity described in any work on Rejang that we have been able to access. Future fieldwork is necessary to provide this confirmation.

As McGinn argues, the agent in such cases are not conventional preverbal subjects, but instead in a lower position and cliticized to the verb. Consider the contrast between the baseline canonical declarative in (23a) and the grammatical active theme relative in (23b), both with the preverbal inchoative auxiliary *mulaé* and the same active verb. The canonical subject position precedes preverbal tense/aspect auxiliaries as in (23a), but pronominal agents may appear in an immediately preverbal position (23b), which allows the active theme to be extracted. We also see here that the first-singular pronoun *uku* appears in a reduced form *ku* when in this immediately preverbal position.

(23) **Subject pronoun positions:** (McGinn, 1998: 373 ex. 38, 372 ex. 36)

a. **Uku** mulaé toton pilem o.
1sg AUX ACT-watch movie the
'I began to watch the movie.'

b. pilem [gi mulaé **ku** toton ___ kelem] o
movie C_{gi} AUX 1sg ACT-watch last.night the
'the movie that I began to watch.'

As noted in McGinn 1989, the structure here is reminiscent of the so-called "Passive type 2" in nearby Malayic languages (see e.g. Chung, 1976; Arka and Manning, 1998), illustrated with Indonesian in (24), where an agent of limited or reduced size (Nomoto, 2021) immediately precedes the lexical verb and which then allows for theme extraction.

(24) **Theme relativization across a pronominal agent in Indonesian:**
Orang [yang sudah **saya** tulis-i ___ surat] adalah paman-mu.
man C_{REL} PERF 1sg write-APPL letter be uncle-2sg
'The man that I wrote a letter was your uncle.' (Chung, 1976: 72 ex. 91b)

However, importantly, the Rejang structures (22, 23b) retain active voice morphology on the verb, whereas no voice morphology appears in the Malayic structures as in (25). See McGinn 1989 for further discussion and comparison of these forms in Rejang and Malayic.

[18] McGinn (1982: 24; 1998: 363–364) proposes that such *wh*-questions are pseudoclefts which involve relativization over the gap, which explains the presence of the relative complementizer. Such pseudocleft strategies are common across Austronesian verb-initial languages (Potsdam, 2009).

The possibility of active theme extraction with agent cliticization in Rejang, as in (22) and (23b), is accounted for straightforwardly on our account. It is well documented that clitic arguments do not intervene for extraction in the same way that a full DP does (McGinnis, 1998; Anagnostopoulou, 2003); presumably, cliticization of an argument into the verb renders it invisible for higher [D]-sensitive probes, or it is skipped as it is not eligible for movement (see also Branan to appear). With the subject being a clitic, the object is now the closest accessible DP to the probe on C.

Next we turn to long-distance relativization. First, we note that in non-extraction contexts, embedded clauses take the complementizer *bawo*, in contrast to *gi* in the extraction examples above. This alternation will be important for the discussion that follows.

(25) **Complement clause with complementizer *bawo*:** (McGinn, 1998: 359 ex. 2a)
Alui m-adea' [$_{CP}$ **bawo** Desi teko ceño']
Alui ACT-say C Desi come late
'Alui said that Desi came late.'

Long-distance subject extraction in Rejang comes in two forms. In one option, the embedded clause has the complementizer *bawo*, with a resumptive pronoun in its subject position (26a). In the second option, the embedded clause is headed by *gi*, with a subject gap (26b). In both cases, the higher clause must appear in the passive.

(26) **Two forms of long-distance subject relatives:** (*ibid.*: 368 ex. 26, 28)
a. tun tuey [gi ___$_{CP}$ n-adea' Alui [$_{CP}$ **bawo si** teko ceño']] o
person old C_{gi} PASS-say Alui C 3sg came late the
'the old person of whom it was said that he/she came late'

b. tun tuey [gi ___$_{CP}$ n-adea' Alui [$_{CP}$ **gi** ___$_{DP}$ teko ceño']] o
person old C_{gi} PASS-say Alui C_{gi} came late the
'the old person of whom it was said that he/she came late'

We first consider these complementizer and gap/resumptive alternations. Both reflect strategies for getting the nominal target for relativization to the edge of the embedded CP. The nominal can be base-generated at the embedded clause edge and bind a local pronoun (26a) or it can move from embedded subject position using the complementizer *gi* (26b).

As proposed by McGinn (1998: 368–369), the entire embedded CP itself is then moved to the higher subject position via passivization of 'say,' although it is then pronounced to the right, leaving the CP gap indicated in (26).[19] Movement of the embedded clause to this higher subject position causes the DP at the embedded clause edge to be the highest nominal in the clause, allowing the relative complementizer *gi* to extract it.

(27) **The closest nominal may be *in* a clausal subject:**

$\checkmark$ [PROBE:REL+D] ... [$_{CP}$ $DP_{2[REL],i}$ ... <DP_2>/pro_i ...] ... [DP_1 ... <CP>

In addition, long-distance relativization may proceed across a higher active verb if the higher verb's agent is a preverbal clitic pronoun. See (28a,b). As per our discussion above, clitic pronouns are skipped, making the highest DP in the embedded clause the closest DP for the higher Ā-probe. In (28a), this is the embedded subject. Example (28b) furthermore shows long-distance theme relativization across multiple clitic pronouns, each complementizer along the way attracting the closest DP, skipping an agent clitic.[20]

(28) **Long-distance relativization over active verb with pronominal subject:**

(McGinn, 1998: 373 ex. 37a,b)

a. tun [gi ku m-adea' [$_{CP}$ gi ___ tanang Pak Lu'ea' kelem]] o
person C_{gi} 1sg ACT-say C_{gi} ACT-visit Mr Headman last.night the
'the person that I said visited the headman last night'

b. filem [gi ko m-adea' [$_{CP}$ gi mulaé ko toton ___ kelem]] o
movie C_{gi} 2sg ACT-say C_{gi} AUX 2sg ACT-watch last.night the
'the movie that you said that you began to watch last night'

Further evidence that the extraction restriction in Rejang reflects [PROBE:Ā+D] on C, and need not target subjects, comes from possessor relatives. Possessor relativization in Rejang is possible with a resumptive pronoun, but generally only for possessors of subjects:

[19] See Rackowski and Richards 2005 and Van Urk and Richards 2015 for similar analyses of long-distance extraction in Tagalog and Dinka, respectively. The postverbal position of the agent *Alui* in (26) forms an argument against passive morphology simply appearing as a reflection of extraction across the verb.

[20] Both our account and the discussion in McGinn 1998 predict the availability of long-distance theme relativization via passivization in both higher and lower clauses, but examples of this form are not discussed there.

(29) **Possessor relativization with resumptive:** (*ibid.*: 370 ex. 33a,b)

a. tun [gi [nyung *(**ne**)] panjang] o
person C_{gi} nose 3sg.GEN long the
'the person whose nose is long'

b. *tun [gi Alui klea' [ngenyan **ne**]] o
person C_{gi} Alui ACT-see wife 3sg.GEN the
Intended: 'the person whose wife Alui saw'

Relativization of object possessors is however possible if the subject is a clitic pronoun, as is attested in the Dibul 2019 translation of the Quran:

(30) **Object possessor relativization over pronominal subject:**
api bae [de[22] Keme m-anjang [omor ne]]
who only C_{gi} 1pl.excl ACT-long age 3sg.GEN
'whomsoever We cause to live long' (Quran 36:68; Shakir translation)
literally 'whoever we lengthen their age'

These examples show that possessor relativization is locality-sensitive and does not reflect a free process of pronominal binding. The use of the *gi* complementizer — which correlates with movement, as we saw in (26) — also supports the view that possessor relativization as in (29a) and (30) involves extraction.

We can imagine two possible analyses for this possessor relativization: one where the target of relativization originates in the possessor position, with its trace pronounced as the pronoun *ne* (as in a recent proposal in Jeoung 2018 and suggested by a reviewer), or with the target of relativization being generated at the edge of the DP and locally binding the possessive pronoun. In either case, the restriction of possessor relativization to possessors of the highest nominal (normally the subject) is explained by our analysis as analogous to the long-distance extraction facts above: the target in these cases is at the edge of the highest DP and therefore counts as "closest" to the probe.

(31) **The closest nominal may be *in* a nominal subject:**
✓[PROBE:REL+D] ... [$_{DP1}$ $DP_{2[REL],i}$ ... N <DP_2>/pro_i] ...

[22] The form of the *gi* complementizer is subject to dialectal variation.

In sum, a close look at the apparent subject-only extraction restriction with the *gi* complementizer in Rejang — and the shape of its various exceptions — provides strong motivation for the theory of probing presented here. Rejang allows for extraction of non-subject nominals across clitic agents with *gi*, and provides strategies to place nominals at the edge of subjects — both nominal and clausal — causing said nominal to become the closest nominal to the probe and licit targets for extraction.

4 Evidence from the Keenan and Comrie 1977 typology

We now turn to the broader typology of extraction asymmetries for further evidence of the possibility of Ā-probing for the closest DP (7). Our discussion will center around Keenan and Comrie's work on the typology of relativization (Keenan and Comrie, 1977, 1979; Comrie and Keenan, 1979; hereafter "K&C"). In this work, K&C claim that there exist languages with relativization strategies which apply specifically to subjects but to no other types of arguments.[23] As noted by Deal (2015: 698–699), such forms of relativization may be candidates for being analyzed as involving Ā-probing for the closest DP.

Revisiting the whole set of subject-only relativization strategies identified in K&C's study, we note that the level of detail provided by K&C on individual languages is generally insufficient to determine whether any of these relativization strategies are best analyzed as involving Ā-probing for the closest DP. Thankfully, however, more detailed subsequent studies exist for some of these languages. We conclude that some but not all of K&C's subject-only strategies provide further evidence for our conjecture. This section thus serves two purposes: First, it highlights a few more case studies — from Arabic, Māori, and Toba Batak — which we take to provide strong supporting evidence for the possibility of Ā-probes being restricted to the closest DP. Second, it serves as a note of caution against taking K&C's results to have already established the necessity of Ā-probing for the closest DP in grammar, thereby underscoring the motivation for our current paper.

In their survey, K&C report eleven languages with relativization strategies that can target subjects but not other nominals: Northeast Aoba/Ambae, Arabic, Kiribati (Gilbertese), Iban (Sea

[23] K&C discuss "strategies" of relativization, of which a particular language may have multiple. Individual strategies are distinguished, for example, by whether they involve gapping, resumptive pronouns, or relative pronouns, or by other distinguishing morphosyntactic characteristics.

Dayak), Javanese, Kera, Malagasy, Māori, Minangkabau, Tagalog, Toba Batak. Of these, only Malagasy and Toba Batak are discussed in any detail in K&C 1977. With the exception of Arabic and Kera (East Chadic), all of these languages are Austronesian.[24]

We first discuss Kiribati (Oceanic; VOS), which K&C describe as utilizing a gap strategy for subject relatives (32a) but a pronoun strategy for object relatives (32b). (This same description applies to Northeast Ambae (also Oceanic), Arabic, and Kera as well; we discuss Arabic below.) Object pronouns appear on the verb with the linker *-i-*.[25]

(32) **Kiribati relative clause data in K&C 1979: 337:**

a. te aine are orea te mane
the woman REL hit the man
'the woman who hit the man'

b. te mane are oro-i-**a** te aine
the man REL hit-TR-3SG the woman
'the man that the woman hit'

At first glance, such a language seems amenable to an analysis utilizing an Ā-probe which targets the closest DP: Ā-extraction is limited to the structurally highest DP, the subject, with relativization of other arguments necessitating the use of resumptive pronouns.

However, this view is challenged by data on long-distance object extraction in Sabel 2013.[26] In (33), fronting the embedded plural object 'Mary and Tien' triggers a third-plural pronoun on the embedded verb 'love,' as expected, but also on the higher verb 'know':

[24] Austronesian languages make up 11 out of 49 languages in K&C's survey (1977: Table 1) and thus may be generally overrepresented in their study. Interestingly, K&C discuss Turkish but treat the SR and NSR forms together as a single strategy; see K&C 1979: 348. There is also a note on subject-only participial relatives in "many European languages (e.g. German, Russian, and Polish)" (K&C 1977: 70); see our fn. 8.

[25] Following Harrison 1978, Sabel 2013 calls *-i-* a transitivity marker and we follow their glossing here. It is possible that the verb *orea* in (32a) includes the third-singular object marker but without the transitivity marker, as suggested by Sabel (2013). In fact, K&C gloss *orea* in (32a) as "hit-3sg" and *oroia* in (32b) as "hit-him." Here we follow Trussel (1979: 140–145) in simply glossing the verb as *orea*.

[26] A reviewer notes that Sabel's long-distance extraction facts do not replicate in Fijian, which otherwise exhibits the same behavior as Kiribati. Presumably, this is a point of cross-linguistic variation.

(33) **Long-distance object movement in Kiribati**: (Sabel, 2013: 18)

Meeiri ao Tien aika ti ata-i-**ia** bwa e tangir-i-**ia** Rui.
Mary and Tien REL.3PL 1PL know-TR-3PL that 3SG love-TR-3PL Rui
'It's Mary and Tien that we know that Rui loves.'

Based on such examples, Sabel argues that the object "pronoun" on Kiribati object extraction verbs is a form of agreement fed by successive-cyclic movement of the object. The difference between subject and object relativization as in (32) thus may simply be a morphological one. Kiribati thus represents an instructive test case where further work on the language casts doubt on taking K&C's reported behavior as a natural candidate for analysis using Ā-probing for the closest DP.

K&C's description of relativization as subject-only in Malagasy and Tagalog also does not immediately necessitate Ā-probing for the closest DP. Both are Philippine-type Austronesian languages explicitly discussed by Aldridge (2004, 2008) as amenable to her analysis for syntactic ergativity reviewed above. However, alternative approaches for these languages exist which do not involve such an Ā-probing restriction; see footnote 8.

In some cases, however, subsequent work on these subject-only relativization strategies identified by K&C has led to forceful arguments for Ā-probing that is limited to the closest DP. Such is the case for Toba Batak (Erlewine, 2018), where relativization is limited to the subject (active agent or passive theme), with relativization of transitive themes fed by passivization. (K&C also gives this same description for Iban, Javanese, and Minangkabau — all Indonesian-type Austronesian languages; see fn. 1.) Although Cole and Hermon 2008 proposed that this restriction reflects the "frozen" nature of the non-subject DP arguments, more recently Erlewine 2018 shows that, under certain circumstances, a subject and non-subject DP can be fronted simultaneously. Erlewine argues that the basic extraction restriction thus must be due to Ā-probing being limited to the closest DP — with options for further probing leading to multiple extraction — rather than due to the general immobility of non-subject DPs as Cole & Hermon propose.

Further work on Māori has also led to an analysis involving Ā-probing for the closest DP. Douglas 2018 discusses the fact that clefting in Māori — built from a kind of headless relative clause — can only target subjects of verbal and prepositional phrase predicates, but not subjects of nominal predicates. He argues that this is due to the Māori cleft construction using Ā-probing for the closest DP; in this predicate-initial language, nominal predicates are higher than their argument. How-

ever, as with Toba Batak, K&C's initial description of Māori as having a subject-only relativization strategy does not by itself force a limited Ā-probing analysis: only with further investigative work could a decisive argument for restricted probing be made.

Finally, we discuss the analysis of relativization in Arabic due to Shlonsky 1992. Arabic allows extraction of subjects but requires pronominal resumption for extraction of all other arguments. Shlonsky proposes that Spec,CP in Palestinian Arabic is an A-position, rather than an Ā-position, and thereby obeys the locality profile of A-movement. Although not in the contemporary probe-goal terms used here, Shlonsky's analysis amounts to proposing that the Ā-probe for relativization must target the closest DP. Similar proposals have since been put forward for new information clefts in French (Belletti, 2015) and topicalization in southern Bantu (Bliss and Storoshenko, 2009; Pietraszko, 2021), which target subjects.[27]

In sum, although the evidence provided by K&C is by itself generally insufficient to motivate the existence of Ā-probing restricted to the closest nominal, stronger arguments for proposals of this form have been subsequently developed for three of K&C's languages — Arabic (Shlonsky, 1992), Māori (Douglas, 2018), Toba Batak (Erlewine, 2018) — all three of which are, again, clearly not ergative. These three case studies thus join our discussion of Turkish and Rejang to together form a compelling reason to take seriously that Ā-probing for the closest DP is a true possibility in grammar, and not only in ergative languages. This evidence then in turn indirectly supports the feasibility of the Aldridge 2004, 2008 approach to syntactic ergativity presented in §2.

5 Variation in probing

We have now established that Ā-probing for the closest DP is indeed a strategy employed by the grammar of non-ergative languages. At the same time, we know that languages also employ relativized Ā-probing which can skip intervening nominals without the matching Ā-feature; see (1). In this section, we turn to the nature of this variation. We show that Ā-probing for the closest DP is not a language-level parameter nor a construction-level parameter. Instead, we argue that this choice of restricted probing is made on individual heads.

[27] We thank Asia Pietraszko (p.c.) for bringing these Bantu works to our attention.

5.1 Haya

We begin with a discussion of relativization in Haya, a Bantu language of the Great Lakes region of Africa. Haya shows us that heads at different positions in a single Ā-construction can differ in their choice to employ Ā-probing for the closest DP or not.

Local relativization in Haya can target both subjects and objects:[28]

(34) **Local relativization in Haya is unrestricted:** (Duranti, 1977: 120 ex. 1, 121 ex. 13)

a. embw' é-y-a-ly' ébitooke
dog REL-it-TAM-eat bananas
'the dog that ate bananas'

b. ebitook' eby' émbwá y-á-lya
bananas REL dog it-TAM-eat
'the bananas that the dog ate'

Duranti 1977 shows that, in long-distance relativization of an object in Haya, the object must be promoted to subject before undergoing further Ā-movement to the final landing site in the matrix clause, as demonstrated through the contrast in (35b,c).

(35) **Long-distance theme relative requires passivization:** (*ibid.*: 129 ex. i–iii)

a. Kato n-a-tekelez' [$_{CP}$ aty' omwaana y-a-bon abashaija
Kato PR-he-thinks C child he-PAST-see men
'Kato thinks that the child has seen the men.'

b. *abashaij [abo Kat' a-li-ku-tekelez' [$_{CP}$ aty omwaaana y-a-bona
men REL Kato he-be-to-think C child he-PAST-see
'the men that Kato thinks the child has seen'

c. abashaij [abo kat' a-li-ku-tekelez' [$_{CP}$ ati ba-a-bon-**w** omwaaana
men REL Kato he-be-to-think C they-PAST-see-PASS child
'the men that Kato thinks have been seen by the child'

We can understand this effect as resulting from intermediate movement being driven by [PROBE:REL+D] on embedded C which must target the closest DP, despite the highest clause of the relative not being

[28] The form of the relative marker in (35–36) varies due to agreement. In addition, in subject relatives (35a), the relative marker prefixes to the verb.

restricted in this way.[29] Promotion of an embedded object to subject position through passivization causes the theme to be the highest nominal in the embedded clause, allowing it to then be probed and thereby extracted.

For the analysis developed here, whether or not the extraction restriction will obtain in a configuration involving C is a function of the lexical items in the context. It is relatively common for embedded complementizers to differ from matrix complementizers in terms of what they attract; consider, for instance, the presence vs. absence of T-to-C movement and *do*-support in standard English matrix and embedded questions. Haya, then, is a language in which the outermost relative complementizer (*abo* in (35)) simply probes for [REL], but the embedded complementizer *ati* must find a [REL] target that is the closest DP. Only extraction out of clauses headed by *ati* exhibit the subject-only restriction, even within a single long-distance extraction chain.

5.2 Late Archaic Chinese

Late Archaic Chinese (LAC) exhibits a number of extraction asymmetries which are attributable to restricted Ā-probing by [PROBE:Ā+D] as described here. At the CP level, Aldridge 2019 shows that only subjects can be Ā-extracted to the clause edge in LAC and therefore proposes that these processes involve Ā-probing for the closest DP. In the interest of space, we will not review this evidence from movement to Spec,CP here. Instead, here, we call attention to a restriction on the behavior of non-subject *wh*-phrases which undergo Ā-movement to a clause-medial position in LAC (Aldridge, 2010).

While the canonical word order of LAC is SVO, Aldridge shows that *wh*-objects in LAC appear preverbally:[30]

(36) **Clause-medial *wh*-fronting:** (Aldridge, 2010: 2 ex. 2b, 7 ex. 12b)

a. Wú *shéi* [qī ___]?
I who deceive
'Who do I deceive?'

b. Gōng *shéi* [yù xiāng ___]?
you who want appoint
'Who do you want to appoint?'

[29] Other Bantu languages display similar subject-only restrictions on all clauses; see in particular Demuth and Harford 1999 and Henderson 2006 for more details on these and other patterns of relativization in Bantu.

As Aldridge (2010, 2019) notes, there are two case frames for ditransitives in LAC: one, where the recipient appears preverbally, itself preceded by a marker, *yǐ*, (37a), and another, where the recipient appears postverbally with a preposition, after the theme (37b).

(37) a. ...*yǐ* REC V THEME b. ...V THEME P REC

From the frame in (37a), the recipient can be extracted (38a). From (37b), the theme can be extracted (38b). Ā-movement of the second DP argument of a ditransitive is not attested.

(38) ***Wh*-movement of first argument of each ditransitive:** (*ibid.*: 23 ex. 52a,b)

a. Kè jiāng *hé* [yǐ __ jiào guǎrén]?
you MOD what YI teach me
'What are you going to teach me?'

b. *Hé* [qiú __ yú mín]?
what ask of people
'What will [you] ask of the people?'

These facts support the view that Ā-probing for the closest DP is a possible option for movement-driving heads in general, and not just with C. In LAC, the head which triggers clause-medial movement (perhaps *v*, following Aldridge 2010, 2019) bears [PROBE:WH+D] restricted to the closest DP.[31]

5.3 Non-DP movement in Tagalog and Rejang

The existence of Ā-extraction constructions that must target the closest DP naturally leads to the question or whether and how non-DP constituents can be Ā-extracted. As we have argued in this section, Ā-probing for the closest DP is not a language-level parameter. A language that utilizes [PROBE:Ā+D] in some configurations may also Ā-move non-DPs, but we predict that this will involve different heads or probes.

Here we return to Tagalog — one of the languages that motivated the idea of Ā-probing for the closest DP — and show that the extraction of DPs and non-DPs indeed behaves differently, exemplified here with *wh*-questions. DP fronting, as in (39a) must target the absolutive argument

[30] We follow Aldridge and other historical Chinese linguistics literature in presenting examples with transcriptions based on modern Mandarin pronunciations of the attested examples.

[31] If the predicate-internal subject hypothesis holds of LAC, the agent may be base-generated in Spec,*v*P as well. [PROBE:WH+D] on *v* cannot attract the agent which is already its specifier, making the first object of the ditransitive count as its closest nominal goal. See Branan to appear.

— here, the object of 'eat' — and is separated from the clause with an *ang* case marker. Non-DP fronting, as in (39b) lacks this *ang* marker. In addition, note that the second-position clitic pronoun *=mo* encliticizes to the verb in (39a) but to the *wh*-phrase itself in (39b). See Hsieh 2020a,b for recent, in-depth work on these non-DP extraction constructions in Tagalog.

(39) **DP vs non-DP *wh*-fronting in Tagalog:** (Henrison Hsieh, p.c.)

a. *Ano* ang k<in>ain =mo __ sa kusina?
what ABS <TR.PFV>eat ERG.2SG OBL kitchen
'What did you eat in the kitchen?'

b. *Saan* =mo k<in>ain ang mangga __?
where ERG.2SG <TR.PFV>eat ABS mango
'Where did you eat a mango?'

Structural differences are also observed between DP and non-DP extraction in Rejang. Recall that relativization over DP arguments must target the closest accessible DP. Such examples allow an optional relative pronoun before the complementizer, as in (40).[32]

(40) **Rejang subject relative with optional relative pronoun:** (McGinn, 1982: 20 ex. 39)
tun [(*api*) di ___ k<en>léa Jon] 'o
man who C_{gi} PASS-see Jon the
'the man that was seen by John'

Relativization in Rejang can also target prepositional objects. In such cases a *wh*-containing PP fronts to a position *following* the complementizer.

(41) **Rejang prepositional object relative:** (*ibid.*: 21 ex. 40i,ii)
tun [{di} [PP *magea api*] {*di} Jon m-lié bukew ___] 'o
man C_{gi} to who Jon ACT-give book the
'the man to whom John gave a book'

[32] These examples come from McGinn 1982 which gives Musi dialect forms, including *di* for the complementizer corresponding to *gi* above.

We can understand prepositional object relatives as in (41) as involving a separate, optional process of PP-fronting described by McGinn (1982: 10); this movement "smuggles" the relative operator to a position above the subject for successful targeting by [PROBE:Ā+D]. Although these relative operators are normally optionally realized as a *wh*-word in Spec,CP (as in (40)), the morphological requirements of the preposition require that it be pronounced overtly next to the preposition in (41).

Such structural differences between DP and non-DP extraction are unsurprising under our account. As Aldridge 2004 and much subsequent work has argued for Tagalog, and as we have argued for Rejang, DP movement takes place via an Ā-probe that must target the closest DP, [PROBE:Ā+D]. As this probe cannot target non-DPs, a separate, second Ā-probe — for instance, simply [PROBE:Ā], potentially on a different functional head — must be involved in the fronting of non-DPs.

An important remaining question is why these separate Ā-probes cannot also target DPs, potentially leading to the extraction of non-closest DPs. One possibility, suggested recently by Erlewine (2018) and Hsieh (2020a) for Toba Batak and Tagalog respectively, is that the the Ā-probe which leads to successful non-DP movement cannot target non-closest DPs because doing so would bleed case licensing on the fronted DP. We refer interested readers to these accounts and the argumentation there.

6 Discussion and conclusion

In this paper, we have argued that Ā-probes can be restricted to target only the closest nominal, as originally proposed in Aldridge 2004 as part of an analysis for syntactic ergativity. We showed that this (possibly conceptually surprising) form of restricted Ā-probing is well attested in non-ergative languages. Probing of this form in many languages gives rise to what at first glance may appear to be a subject-only extraction restriction, but one which may be obviated by processes that rearrange nominals as well as subextraction from highest nominals, allowing certain non-subjects to be extracted. We presented examples of extraction restrictions of this form in Turkish and Rejang in detail, and refer the reader to other work describing facts of this form in Arabic, Māori, and Toba Batak — all non-ergative languages.

Establishing the existence of Ā-probing restricted to the closest DP contributes to the growing literature on possible interactions between A- and Ā-features in probe specifications (e.g. Van Urk,

2015; Van Urk and Richards, 2015; Baier, 2018; Erlewine, 2018; Bossi and Diercks, 2019; Colley and Privoznov, 2020; Coon, Baier, and Levin, 2021; Scott, 2021), and the space of variation therein. We furthermore clarified that Ā-probing for the closest DP is a property of specific probes on heads, rather than a language- or construction-level parameter. In particular, languages with constructions that involve Ā-probing for the closest DP often also have other strategies for Ā-extraction that are not so restricted.[33] What this implies, then, is that these grammars must also have mechanisms for choosing between different extraction strategies. Here we pointed to two such mechanisms here: the ability of some complex probes to prefer full matches but optionally allow partial matches (see discussion before section 3.1) and for certain movements to bleed Case licensing and therefore unable to target DPs (section 5.3).

[33] Further support for this view comes from the observation that, even amongst DP arguments, many languages have some Ā-extractions that are more restricted than others in what arguments they can target. See for example discussions of differences between relativization and *wh*-movement in Chukchi (Paleo-Siberian) in Polinsky 1992, 2016 and between topicalization and focus/*wh*-movement in Bikol (Philippine) in Erlewine and Lim 2019, as well as between various Ā-constructions in Kaqchikel (Mayan) in Heaton 2017 ch. 13.

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. 2010. Clause-internal *wh*-movement in archaic chinese. *Journal of East Asian Linguistics* 19:1–36.

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

Aldridge, Edith. 2019. Subject/non-subject movement asymmetries in Late Archaic Chinese. *Glossa* 4:1–38.

Anagnostopoulou, Elena. 2003. *The syntax of ditransitives: Evidence from clitics*. Walter de Gruyter.

Arka, Wayan, and Christopher D. Manning. 1998. Voice and grammatical relations in Indonesian: A new perspective. In *Proceedings of the LFG98 Conference*, ed. Miriam Butt and Tracy Holloway King. CSLI Publications.

Assmann, Anke, Doreen Georgi, Fabian Heck, Gereon Müller, and Philipp Weisser. 2015. Ergatives move too early: on an instance of opacity in syntax. *Syntax* 18:343–387.

Baier, Nico. 2018. Anti-agreement. Doctoral Dissertation, University of California Berkeley.

Belletti, Adriana. 2015. The focus map of clefts: Extraposition and predication. In *Beyond functional sequence*, ed. Ur Shlonsky, 42–59. Oxford University Press.

Belletti, Adriana, and Chris Collins. 2021. Introduction. In *Smuggling in syntax*, ed. Adriana Belletti and Chris Collins, 1–12. Oxford University Press.

Bittner, Maria, and Kenneth Hale. 1996. Ergativity: towards a theory of a heterogeneous class. *Linguistic Inquiry* 27:531–604.

Bliss, Heather, and Dennis Ryan Storoshenko. 2009. Grammaticalization of topic in Shona. Presented at the 3rd International Conference on Bantu Linguistics.

Blust, Robert. 2010. Review: The Austronesian languages of Asia and Madagascar by Alexander Adelaar and Nikolaus P. Himmelmann. *Oceanic Linguistics* 49:302–312.

Blust, Robert. 2018. In memoriam: Richard Bernard McGinn, Jr., 1939–2018. *Oceanic Linguistics* 57:496–502.

Bossi, Madeline, and Michael Diercks. 2019. V1 in Kipsigis: Head movement and discourse-based

scrambling. *Glossa* 4:1–43.

Branan, Kenyon. 2018. Attraction at a distance: $\overline{\text{A}}$-movement and case. *Linguistic Inquiry* 49:409–440.

Branan, Kenyon. to appear. Locality and anti-locality: The logic of conflicting requirements. *Linguistic Inquiry* .

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

Cagri, Ilhan Merih. 2005. Minimality and Turkish relative clauses. Doctoral Dissertation, University of Maryland.

Cagri, Ilhan Merih. 2009. Arguing against subject incorporation in Turkish relative clauses. *Lingua* 119:359–373.

Campana, Mark. 1992. A movement theory of ergativity. Doctoral Dissertation, McGill.

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. 1991. Some notes on economy of derivation and representation. In *Principles and parameters in comparative grammar*, ed. Robert Freidin, 417–454. Cambridge, Mass.: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT 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.

Chung, Sandra. 1976. An object-creating rule in Bahasa Indonesia. *Linguistic Inquiry* 7:41–87.

Chung, Sandra. 1982. Unbounded dependencies in Chamorro grammar. *Linguistic Inquiry* 13:39–78.

Chung, Sandra. 1994. *Wh*-agreement and "referentiality" in Chamorro. *Linguistic Inquiry* 25:1–44.

Chung, Sandra. 1998. *The design of agreement: Evidence from Chamorro*. University of Chicago Press.

Cole, Peter, and Gabriella Hermon. 2008. VP raising in a VOS language. *Syntax* 11:144–197.

Colley, Justin, and Dmitry Privoznov. 2020. On the topic of subjects: Composite probes in Khanty. In *Proceedings of NELS 50*, ed. Mariam Asatryan, Yixiao Song, and Ayana Whitmal, 111–124.

Collins, Chris. 2001. Economy conditions in syntax. *The Handbook of Contemporary Syntactic Theory* 45–61.

Comrie, Bernard, and Edward L. Keenan. 1979. Noun phrase accessibility revisited. *Language* 55:649–664.

Coon, Jessica, Nico Baier, and Theodore Levin. 2021. Mayan Agent Focus and the ergative extraction constraint: Facts and fictions revisited. *Language* 97:269–332.

Coon, Jessica, Pedro Mateo Pedro, and Omer Preminger. 2014. The role of case in A-bar extraction asymmetries: Evidence from Mayan. *Linguistic Variation* 14:179–242.

Deal, Amy Rose. 2015. Ergativity. In *Syntax: Theory and analysis. an international handbook*, ed. Tibor Kiss and Artemis Alexiadou, 654–707. de Gruyter.

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.

Demuth, Katherine, and Carolyn Harford. 1999. Verb raising and subject inversion in Bantu relatives. *Journal of African Languages and Linguistics* 20:41–62.

Dibul, Ahmad. 2019. *Al Quran ngen tai ne lem Baso Jang* [*The Quran with its translation in Bahasa Rejang*]. Institut Agama Islam Negeri (IAIN) Curup [State College for Islamic Studies of Curup].

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

Duranti, Alessandro. 1977. Relative clauses. In *Haya grammatical structures*, ed. Ernest Rugwa Byarushengo and Alessandro Duranti, 119–132. University of Southern California.

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

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, and Cheryl Lim. 2019. Bikol clefts and topics and the Austronesian extraction restriction. Manuscript, National University of Singapore.

Georgopoulos, Carol. 1985. Variables in Palauan syntax. *Natural Language & Linguistic Theory* 3:59–94.

Georgopoulos, Carol. 1991. *Syntactic variables: Resumptive pronouns and A' binding in Palauan*. Springer.

Hankamer, Jorge, and Laura Knecht. 1976. The role of the subject/non-subject distinction in determining the choice of relative clause participle in Turkish. In *Proceedings of NELS 6*, volume 2, 123–135.

Harrison, Shelley P. 1978. Transitive marking in Micronesian languages. In *Proceedings of ICAL 2*, 1067–1127.

Heaton, Raina. 2017. A typology of antipassives, with special reference to Mayan. Doctoral Dissertation, University of Hawai'i.

Henderson, Brent Mykel. 2006. The syntax and typology of Bantu relative clauses. Doctoral Dissertation, University of Illinois at Urbana-Champaign.

Himmelmann, Nikolaus P. 2002. Voice in western Austronesian: an update. In *The history and typology of western Austronesian voice systems*, ed. Fay Wouk and Malcolm Ross, 7–16. Canberra: Pacific Linguistics.

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

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

Hsieh, Henrison. 2020b. On the structure of Tagalog non-DP extraction. In *Proceedings of AFLA 26*, ed. Ileana Paul, 124–141.

Jeoung, Helen. 2018. Possessors move through the edge, too. *Glossa* 3:1–35.

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

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

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

Keenan, Edward L., and Bernard Comrie. 1979. Data on the noun phrase accessibility hierarchy.

Language 55:333–351.

Kelepir, Meltem. 2001. Topics in Turkish syntax: Clausal structure and scope. Doctoral Dissertation, Massachusetts Institute of Technology.

Kornfilt, Jaklin. 1984. Case marking, agreement, and empty categories in Turkish. Doctoral Dissertation, Harvard.

Kornfilt, Jaklin. 1997. *Turkish*. Routledge.

Kornfilt, Jaklin. 2000. Some syntactic and morphological properties of relative clauses in Turkish. In *The syntax of relative clauses*, ed. Artemis Alexiadou, Paul Law, André Meinunger, and Chris Wilder, 121–160. John Benjamins.

Kornfilt, Jaklin. 2008. Subject case and Agr in two types of Turkish RCs. In *Proceedings of WAFL 4*, ed. S. Ulutaş and Cedric Boeckx, 145–168. MIT Working Papers in Linguistics.

Legate, Julie Anne. 2002. Warlpiri: Theoretical implications. Doctoral Dissertation, Massachusetts Institute of Technology.

Legate, Julie Anne. 2008. Morphological and abstract case. *Linguistic Inquiry* 39:55–101.

Levin, Theodore. 2018. K'ichean syntactic ergativity: Locality and intervention in A'-movement. Manuscript, National University of Singapore.

Martinović, Martina. 2015. Feature geometry and head-splitting: Evidence from the morphosyntax of the Wolof clausal periphery. Doctoral Dissertation, University of Chicago.

McGinn, Richard. 1982. *Outline of Rejang syntax*, volume 14 of *NUSA: Linguistic Studies of the Languages of Indonesia*.

McGinn, Richard. 1989. The animacy hierarchy and Western Austronesian languages. In *Proceedings of ESCOL 6*, 207–217.

McGinn, Richard. 1998. Anti-ECP effects in the Rejang language of Sumatra. *Canadian Journal of Linguistics* 43:359–376.

McGinnis, Martha Jo. 1998. Locality in A-movement. Doctoral Dissertation, Massachusetts Institute of Technology.

Murasugi, Kumiko. 1992. Crossing and nesting paths: NP movement in accusative and ergative languages. Doctoral Dissertation, Massachusetts Institute of Technology.

Nakamura, Masanori. 1996. Economy of chain formation. Doctoral Dissertation, McGill.

Nie, Yining. 2019. Raising applicatives and possessors in Tagalog. *Glossa* 4:1–30.

Nomoto, Hiroki. 2021. Bare passive agent hierarchy. In *Proceedings of AFLA 27*, ed. Henrison Hsieh and Keely New, 57–70.

Ordóñez, Francisco. 1995. The antipassive in Jacaltec: A last resort strategy. *Catalan Working Papers in Linguistics* 4:329–343.

Pearson, Matthew. 2001. The clause structure of Malagasy: A Minimalist approach. Doctoral Dissertation, University of California at Los Angeles.

Pearson, Matthew. 2005. The Malagasy subject/topic as an A′-element. *Natural Language & Linguistic Theory* 23:381–457.

Pesetsky, David. 2017. Complementizer-trace effects. In *Blackwell Companion to Syntax*, ed. Martin Everaert and Henk van Riemsdijk. Wiley-Blackwell, second edition.

Pesetsky, David, and Esther Torrego. 2001. T-to-C movement: Causes and consequences. In *Ken Hale: A life in language*, ed. Michael Kenstowicz, 355–425. MIT Press.

Pietraszko, Asia. 2021. The coming apart of case and focus in Bantu. *Natural Language & Linguistic Theory* 39:579–599.

Polinsky, Maria. 1992. Relativization in Chukchi. In *The non-Slavic languages of the USSR*, 241–262. Chicago Linguistic Society.

Polinsky, Maria. 2016. *Deconstructing ergativity: Two types of ergative languages and their features*. Oxford University Press.

Polinsky, Maria. 2017. Syntactic ergativity. In *Blackwell Companion to Syntax*, ed. Martin Everaert and Henk van Riemsdijk. Wiley-Blackwell, second edition.

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.

Rizzi, Luigi. 1990. *Relativized minimality*. MIT Press.

Ross, Malcolm. 2002. The history and transitivity of western Austronesian voice and voice marking. In *The history and typology of western Austronesian voice systems*, ed. Fay Wouk and Malcolm Ross, 17–62. Canberra: Pacific Linguistics.

Sabel, Joachim. 2013. Configurationality, successive cyclic movement, and object agreement in Kiribati and Fijian. *Linguistische Berichte* 233:3–22.

Sağ-Parvardeh, Yağmur. 2019. The semantics of number marking: Reference to kinds, counting, and optional classifiers. Doctoral Dissertation, Rutgers.

Scott, Tessa. 2021. Formalizing two types of mixed A/Ā movement. Manuscript, University of California at Berkeley.

Sheehan, Michelle. 2017. Parameterizing ergativity: An inherent case approach. In *The Oxford handbook of Ergativity*, ed. Jessica Coon, Diane Massam, and Lisa Travis, 59–85. Oxford University Press.

Shlonsky, Ur. 1992. Resumptive pronouns as a last resort. *Linguistic Inquiry* 23:443–468.

Starke, Michal. 2001. Move dissolves into Merge: a theory of locality. Doctoral Dissertation, University of Geneva.

Taylan, Eser Erguvanlı. 1984. *The function of word order in Turkish grammar*. University of California Press.

Temürcü, Ceyhan. 2001. Word order variations in Turkish: Evidence from binding and scope. Master's thesis, Middle East Technical University.

Trussel, Stephen. 1979. *Kiribati (Gilbertese) grammar handbook*. Peace Corps Language Handbook Series.

Underhill, Robert. 1972. Turkish participles. *Linguistic Inquiry* 3:87–99.

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

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

Woolford, Ellen. 1997. Four-way case systems: Ergative, nominative, objective and accusative. *Natural Language & Linguistic Theory* 15:181–227.

Woolford, Ellen. 2006. Lexical case, inherent case, and argument structure. *Linguistic Inquiry* 37:111–130.

Öztürk, Balkız. 2009. Incorporating agents. *Lingua* 119:334–358.