

This printout has been approved by me, the author. Any mistakes in this printout will not be fixed by the publisher. Here is my signature and the date _____.

Raising to Object out of CP as Embedded Left Dislocations: Evidence from Three Formosan Languages

Victoria Chen and Shin Fukuda

1. Introduction

Raising to Object (RTO) refers to a phenomenon in which a semantically and thematically embedded element appears to surface outside of the embedded clause and exhibits characteristics typical of matrix objects, in terms of case marking, agreement, and binding. Studies have shown that putative RTO constructions in different languages exhibit considerable cross-linguistic variations with respect to (i) the structure of the embedded clause, (ii) the actual position of the “raised” phrase (henceforth XP), (iii) restrictions on the phrase inside the embedded clause that is co-referential with XPs (henceforth YPs), and (iv) how the relationship between XP and YP is established (e.g., Massam 1985; Bruening 2001; Polinsky and Potsdam 2001; Alboiu and Hill 2013). This paper analyzes apparent cases of RTO in Puyuma (1a), Amis (1b), and Seediq (1c), three Formosan languages that belong to three primary branches of the Austronesian family and exhibit highly productive RTO constructions. To the best of our knowledge, this is the first study to examine RTO in Formosan languages.

- | | | | | | | | | | |
|-----|----|----------------------------------|---|-----------------------|---------------------|----------|---------------|---|--------|
| (1) | a. | ma-ladram=ku
AV-know=1SG.ABS | kan
<i>SG.OBL A_i</i> | Atrungi
[C | [dra m-uka
AV-go | i
LOC | Arasip
A | <i>eci</i> . ¹
<i>eci</i> | PUYUMA |
| | b. | ma-fana' kaku
AV-know 1SG.ABS | ci
<i>PN L_r-OBL</i> | Lisin-an
[C | [Ø tayra
AV-go | i
LOC | Kalingku
K | <i>eci</i> .
<i>eci</i> | AMIS |
| | c. | kela-un=mu
know-PV=1SG.ERG | ka
<i>ABS I_i</i> | Ikungi
[C | [Ø m-usa
AV-go | Ø
LOC | Skangki
S | <i>eci</i> .
<i>eci</i> | SEEDIQ |

As will be shown in this paper, despite the superficial similarities evident in (1), RTO in these three Formosan languages diverge in two important regards. First, RTO in Amis and Seediq requires the YP inside the embedded clause to be an absolutive (ABS)-marked phrase, while such a restriction is absent in Puyuma RTO. Second, the XP-YP relation obeys islands in Seediq RTO, while the same relation is island-*insensitive* in Amis and Puyuma RTO. An obvious question is where these variations come from.

In this paper, we argue that the similarities and differences among RTO in these three languages can be accounted for if they are analyzed as instances of embedded left-dislocation that utilize independently motivated strategies to establish the relation between the left-dislocated phrase (the XP) and the CP. Following Landau (2011), we argue that in RTO, a CP and an XP may establish their relation in three different ways. First, the relation between a saturated (i.e., propositional) CP and an XP may be established pragmatically through *the aboutness condition* (2a), which is what we propose to be the case in Puyuma RTO. Alternatively, an XP-CP relation may be established via co-indexation

*Victoria Chen, University of Hawai'i at Mānoa, ychen6@hawaii.edu. Shin Fukuda, University of Hawai'i at Mānoa, fukudash@hawaii.edu. This project is supported by funding from the Department of Linguistics at University of Hawai'i and the Linguistics Institute of Academia Sinica. We thank Robert Blust, Kamil Deen, Ying-chin Lin, Diane Massam, William O'Grady, Yuko Otsuka, Masha Polinsky, and the audience at WCCFL 33 and AFLA 22 for comments and suggestions on this study. Our sincere thanks to our consultants: Min-ying Sun, Chung-yi Tsai [*Puyuma*], Chin-hua Wu-Tseng, Wan-sung Lin [*Amis*], Yu-ju Chu, Chiu-kui Tai, and Yu-hsia Lin [*Seediq*].

¹Abbreviations: ABS=absolutive, AGR=agreement, AGT=agent, AV=actor voice, C=complementizer, DF=definite, ERG=ergative, ID=indefinite, IRR=irrealis, LK=linker, LOC=locative, LV=Locative voice, NEG=negation, NMZ=nominalizer, OBL=oblique, PERF=perfective, PL=plural, POSS=possessive, PV=patient voice, SG=singular.

of the XP and a null operator (Op) in [Spec, CP] (e.g., Anagnostopoulou 1997). As argued by Landau (2011), CPs with an operator (i.e., “predicative CPs” in Landau’s term) can be divided into two types: one that involves a base-generated Op (2b), and the other that involves an A’-moved Op (2c). We argue that Amis RTO involves a base-generated Op (2b), while Seediq RTO involves an Op movement (2c).

- (2) a. V_{MATRIX} XP + [CP $\quad \quad \quad$] via the aboutness condition [Puyuma RTO]
 b. V_{MATRIX} XP_i [CP Op_i *pronoun_i*] [Amis RTO]
 c. V_{MATRIX} XP_i [CP Op_i <*t_i*>] [Seediq RTO]

If the proposed analysis is on the right track, the data from three Formosan languages reveal that closely related languages may utilize slightly different mechanisms that are independently available to realize what can be descriptively characterized as RTO constructions. Importantly, however, RTO in none of the three languages involves actual movement of the XP out of the CP, thus avoiding violations of the Phase Impenetrability Condition and improper movement configuration.

In what follows, we first present the core data in Section 2. In Section 3, we propose our analysis of the basic structure of RTO across the three languages as involving a finite CP with a base-generated left-dislocated XP that adjoins to it. Section 4 presents our account for the differences among RTO constructions in these languages, as summarized in (2). Section 5 discusses the implications of the proposed analysis for RTO cross-linguistically, and concludes the paper.

2. Similarities and Differences in RTO in Puyuma, Amis and Seediq

Puyuma, Amis, and Seediq are three Philippine-type Formosan languages that belong to three different Austronesian primary branches (Blust 1999). They are all predicate-initial, and exhibit what can be described as an *Absolutive as unique* constraint, i.e., every clause must possess one and only one absolutive-marked phrase. Table 1 summarizes the mapping between case markers and different thematic roles of arguments in the RTO data to be discussed in this paper.

	Agent	Patient/Theme	Location
<i>Actor voice</i> (AV)	[ABS]	[(OBL)]	[(OBL)]
<i>Patient voice</i> (PV)	[ERG]	[ABS]	—
<i>Locative voice</i> (LV)	[ERG]	[ABS]	—

Table 1: Case pattern in RTOs across Puyuma, Amis & Seediq

2.1. Similarities

2.1.1. Fully productive RTO with finite CP complements

RTO across the three Formosan languages is associated with knowledge/perception verbs (e.g., *know, see, dream, hear, forget, fear*) and finite CP complements. The finite CP analysis is built on the observation that the embedded clauses of RTO are fully compatible with different aspect markers and voice markers, unlike infinitives (control and restructuring complements) in the same languages, which disallow aspect markers and voice markers other than AV. Additionally, finite CPs and RTO complements in Puyuma carry an overt complementizer, which never appears inside infinitives (Table 2). According to our fieldwork, RTO in all three languages is fully productive and compatible with all CP-taking verbs.

	Finite CPs/RTO complements	Control/Restructuring infinitives
<i>Aspect markers</i>	✓	×
<i>Unrestricted voice type</i>	✓	× (AV-only)
<i>*Overt complementizer</i>	✓*Puyuma	×

Table 2: Morphosyntactic constraints in types of complementation across the three languages

2.1.2. The matrix behavior of the XP

In all three languages, XP behaves like a matrix object in terms of case-licensing and reflexive binding. The case status of an XP is always determined by the appropriate matrix case licensor and presents no case-connectedness effect, as shown in (3a)-(3c). The XP is thematically identified with the absolutive argument of the embedded clause, yet bears Oblique (OBL) case under a matrix AV verb.

- (3) a. ma-ladram=ku *kana* *ngiawi* [dra karatr *eci*(ABS) kana suan].
 AV-know=1SG.ABS **DF.OBL** *cat* [C <AV>bite *eci*(ABS) DF.OBL dog]
 ‘I know (about the cat) that it bit the dog.’ PUYUMA
- b. ma-fana’ kaku *ci* *Sawmah-ani* [Ø mi-sakilif *eci*(ABS) ci Kulas-an].
 AV-know 1SG.ABS **PN S-OBL** [C AV-lie *eci*(ABS) PN K-OBL]
 ‘I know (about Sawmah) that she lied to Kulas.’ AMIS
- c. me-’isug=ku *Imin-Ø* [Ø s<m>ipaq huling=mu *eci*(ABS)].
 AV-fear=1SG.ABS **I-OBL** [C <AV>beat dog=1SG.POSS.(OBL) *eci*(ABS)]
 ‘I fear (about Imin) that she will beat my dog.’ SEEDIQ

In addition, when XPs are anaphors, they are obligatorily bound by the matrix external argument. In RTO sentences with a reflexive XP (4)-(6), the reflexivization relation can only be bound by the matrix subject ((4b), (5b), and (6b)) and never by the embedded subject. As a result, RTO with a reflexive XP always receives a different interpretation from its non-raising counterparts ((4a), (5a), and (6a)).

- (4) a. ma-tiya i Kisaw [dra tu=satra’-ay kan Siber kantaaw]. PUYUMA
 AV-dream SG.ABS K [C 3.ERG=slap-LV SG.ERG S himself.ABS]
 ‘Kisaw_i dreamt that [Siber_k slapped himself_{i/k}].’
- b. ma-tiya i Kisaw *taytaaw* [dra tu=satra’-ay kan Siber *eci*].
 AV-dream SG.ABS K **himself.OBL** [C 3.ERG=slap-LV SG.ERG S *eci*]
 ‘Kisaw_i dreamt that [Siber_k slapped him_{i/*k}].’
- (5) a. ma-lemed ni Kulas [Ø ma-palu ni Mayaw cingra]. AMIS
 PV-dream ERG K [C PV-beat ERG M himself.ABS]
 ‘Kulas_i dreamt that [Mayaw_k beat himself_{i/k}].’
- b. ma-lemed ni Kulas *cingra* [Ø ma-palu ni Mayaw *eci*].
 PV-dream ERG K **himself.ABS** [C PV-beat ERG M *eci*]
 ‘Kulas_i dreamt that [Mayaw_k beat him_{i/*k}].’
- (6) a. spi-an Ø Walisi [Ø s<n>ipaq Watan_k ka heya nanak_k]. SEEDIQ
 dream-LV ERG W [C <PV>beat Watan.ERG ABS himself_k]
 ‘Walis_i dreamed that Watan_k slapped himself_{i/k}.’
- b. spi-an Ø Walisi *ka* *heya nanaki* [Ø s<n>ipaq Ø Watan_k *eci*].
 dream-LV ERG W **ABS** **himselfi** [C <PV>beat ERG W *eci*]
 ‘Walis_i dreamed that [Watan_k slapped him_{i/*k}].’

2.1.3. The absence of reconstruction effects

In all three languages, the XP in RTO shows no reconstruction effect. As shown below (7)-(9), when a simple CP complement involves a quantifier subject and a possessive direct object, a bound variable reading of the possessive direct object obtains. Therefore, in (7a), the referent of *kantu walak* ‘3.POSS.OBL child’ depends on the referent of its binder *taynaynayan* ‘all mothers’. In contrast, if the possessive direct object is “raised” and becomes an XP (7b), the bound variable reading is no longer available. Hence, the possessive XP can only receive a specific interpretation. Under (7), for instance, the XP can only refer to a specific set of children, i.e., ‘I know that all mothers_{i/*k} love their_{i/k} children’.

- (7) a. ma-ladram=ku [dra sagar kantu walak na taynaynayan]. PUYUMA
 AV-know=1SG.ABS [C like.AV 3.POSS.OBL child DF.ABS all.mothers]
 ‘I know that all mothers_i love their_{i/*k} children.’
- b. ma-ladram=ku *kantu* *walak* [dra sagar *eci* na taynaynayan].
 AV-know=1SG.ABS **3.POSS.OBL** **child** [C like.AV *eci* DF.ABS all.mothers]
 ‘I know that all mothers_i love their_{i/*k} children.’
- (8) a. ma-lemed aku [Ø maemin ma-palu nuna wawa ku wacu nira]. AMIS
 PV-dream 1SG.ERG [C all PV-beat that.ERG child ABS dog 3SG.POSS]
 ‘I dream that all children_i beat their_{i/*k} dogs.’
- b. ma-lemed aku *ku* *wacu* *nira* [Ø maemin ma-palu nuna wawa *eci*].
 PV-dream 1SG.ERG **ABS** **dog** **3SG.POSS** [C all PV-beat that.ERG child *eci*]
 ‘I dream that all children_i beat their_{i/*k} dogs.’
- (9) a. kela-un=mu [Ø qelu-un bi de-bubu ka laqi=deha]. SEEDIQ
 know-PV=1SG.ERG [C love-PV very all-mothers ABS child=3PL.POSS]
 ‘I know that all mothers_i love their_{i/*k} children.’
- b. kela-un=mu *ka* *laqi=deha* [Ø qelu-un bi de-bubu *eci*].

know-PV=1SG.ERG ABS *child*=3PL.POSS [C love-PV very all-mothers *eci*]
 ‘I know that all mothers love their children*_{i/k}.’

2.2. Differences

Despite the similarities described above, RTO constructions in the three languages diverge in two important regards: (i) restriction on the YP and (ii) sensitivity to islands.

2.2.1. Amis/Seediq vs. Puyuma: The ABS-only constraint on the YP

First, Puyuma RTO differs from that in Amis and Seediq in that RTO in the latter two languages requires the XP to be identified with the embedded absolutive argument, while such a restriction is absent in Puyuma RTO. As shown in (10a) and (11a), Puyuma RTO allows the XP to be identified with an embedded ergative or oblique argument as long as it is definite, while in Amis and Seediq RTO ((10b-c) and (11b-c)), an XP identified with any non-absolutive element results in ungrammaticality.

- (10) YP as an embedded oblique
- a. ma-ladram=ku *i* *Arasipi* [dra m-uka *eci* i Atrung]. (PUYUMA)
 AV-know=1SG.ABS LOC *A_i* [C AV-go *eci* SG.ABS Atrung]
 ‘I know (about Arasip) that Atrung went there.’
- b. *ma-fana’ kaku *i* *Kalingku* [Ø tayra *eci* Ø-ci Lisin]. (AMIS)
 AV-know 1SG.ABS LOC *K_i* [C AV-go *eci* ABS-PN Lisin]
 ‘(I know (about Kalingku) that Lisin went there.)’
- c. *kela-un=mu Ø *Skangki* [Ø m-usa *eci* ka Ikung]. (SEEDIQ)
 know-LV=1SG.ERG LOC *S_i* [C AV-go *eci* ABS Ikung]
 ‘(I know (about Skangki) that Ikung went there.)’
- (11) YP as an embedded ergative
- a. ma-ladram=ku *kan* *Isawi* [dra tu=trakaw-aw na ngiaw *eci*]. (PUYUMA)
 AV-know=1SG.ABS SG.OBL *I_i* [C 3.ERG=steal-PV DF.ABS cat *eci*]
 ‘I know (about Isaw) that he stole your cat.’
- b. *ma-fana’ kaku *ci* *Sawmah-an* [Ø ma-keter *eci* kuna wacu]. (AMIS)
 AV-know 1SG.ABS PN S-OBL *i* [C PV-scold *eci* that.ABS dog]
 ‘(I know (about Sawmah) that she scolded that dog.)’
- c. *kela-un=mu *ka* *Hubi* [Ø ’uq-un *eci* ka sari=su]. (SEEDIQ)
 know-PV=1SG.ERG ABS *H_i* [C eat-PV *eci* ABS taro=2SG.POSS]
 ‘(I know (about Hubi) that she ate your taro.)’

2.2.2. Seediq vs. Amis/Puuma: The island-sensitivity of the XP-YP relation

Seediq RTO is distinct from RTO in the other two languages regarding its XP-YP relation’s sensitivity to islands. In Seediq, the XP-YP relation respects both complex NP and adjunct islands ((12a) and (13a)), while the same relation in Puyuma and Amis RTO is immune to these islands ((12b-c) and (13b-c)).

- (12) Complex NP islands
- a. kilengaw=ku *kan* *Isawi* [dra ma-ladram=yu [kana kasaerueru [dra sagar *eci* dra le’u]]]. (PUYUMA)
 AV.hear=1SG.ABS SG.OBL *I_i* [C AV-know=2SG.ABS [OBL anecdote [C like.AV *eci* OBL owl]]]
 ‘I heard that you know the anecdote that Isaw likes owls.’
- b. ma-tengil aku Ø-ci *Kulas* [Ø ma-fana’ kisu [a califacif [Ø ma-ulah *eci* takuwanan]]]. (AMIS)
 PV-hear 1SG.ERG ABS-PN *K* [C AV-know 2SG.ABS [LK rumor [C AV-like *eci* 1SG.OBL]]]
 ‘I heard that you know the rumor that Kulas has a crush on me.’
- c. *q<um>bahang=ku *Hubi* [Ø kela-un=su [ka kari shelisum [Ø qiyut *eci* babuy-Ø]]]. (SEEDIQ)
 <AV>hear=1SG.ABS *H_i*-OBL [C know-PV=2SG.ERG [ABS anecdote [C <AV>bite *eci* pig-OBL]]]
 ‘(I heard that you know the anecdote that Hubi bit pigs.)’
- (13) Adjunct islands
- a. ma-ladram=ku *kan* *Siber* [dra ka-ilemus=ku [an tu=pukpuk-aw=yu *eci*]]. (PUYUMA)
 AV-know=1SG.ABS SG.OBL *S_i* [C IRR.AV-be.angry=1SG.ABS[if 3.ERG=beat-PV=2SG.ABS *eci*]]]
 ‘I know (about Siber) that I will be angry if he beats you.’
- b. ma-fana’ Ø-ci *Kulas* *tuna* *wacu* [Ø t<um>angic kaku [anu ma-patay *eci*]]. (AMIS)
 AV-know ABS-PN *K* OBL.that dog [C <AV>cry 1SG.ABS [if AV-die *eci*]]]

- ‘Kulas knows (about the dog) that I will cry if it dies.’
- c. *kela-un=mu *ka Imini* [ya’asa m-huqil ka huling [kika lingis *eci*]].
 know-PV=1SG.ERG *ABS I_i* [because AV-die ABS dog [so <AV>cry *eci*]]
 (‘I heard (about Imin) that because the dog died, she is crying.)

SEEDIQ

2.3. Section Summary

In sum, RTO constructions in Puyuma, Amis, and Seediq are similar in terms of (i) the finite CP status of their complements, and (ii) the lack of reconstruction effects with their XPs, while they diverge in terms of the case restriction on the YP and the XP-YP relation’s sensitivity to islands. Puyuma RTO is the most unrestricted, imposing no restriction on YP or the XP-YP relation. Amis RTO obeys the absolutive-only constraint on the YP, but shows no island-sensitivity of the XP-YP relation. Seediq RTO is the most restricted, as it is subject to both the absolutive-only constraint on the YP and island conditions. Table 3 summarizes the empirical observations presented so far.

	PUYUMA	AMIS	SEEDIQ
<i>Reconstruction effects</i>	×	×	×
<i>ABS-restriction on the YP</i>	×	✓	✓
<i>Island-sensitivity</i>	×	×	✓

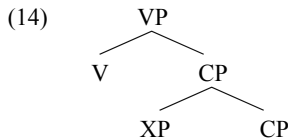
Table 3: Similarities and differences in RTO across the three languages

3. Formosan RTO as a Case of Embedded Left-Dislocation

Based on the empirical observations discussed above, we present our analysis of the basic structure of RTO, according to which RTO constructions are instances of embedded left-dislocation. We first motivate the left-dislocated adjunct analysis for the XPs based on the similarities between XPs and base-generated left-dislocated phrases. We then present our account of the matrix object-like behavior of XPs.

3.1. Formosan RTO Involves a Finite CP with a Base-Generated Left-Dislocated Phrase

In the previous discussion (Section 2.1.1), we showed that RTO in the three languages involves a finite CP complement, based on the unrestricted aspect and voice marking and overt complementizer observed in the RTO complements. As for XPs, we propose that they are base-generated left-dislocation phrases adjoined to the embedded CPs, as in (14).



This analysis is motivated by the observation that XPs in Formosan RTOs share a number of similarities with base-generated left-dislocated phrases cross-linguistically, such as *left-dislocated phrases* in Niuean (Massam 1985) and Italian (Cinque 1990), *external topics* (as opposed to internal topics) in Mayan languages (Aissen 1992), and *hanging topics* in Greek (Anagnostopoulou 1997). As discussed earlier, XPs in Formosan RTO (i) lack reconstruction effects (2.1.3), and (ii) present no case connectedness effects (2.1.2). Both characteristics are typical of hanging/external topics cross-linguistically. RTO in Puyuma further provides an additional motivation for the base-generated left-dislocated adjunct analysis of XPs. As shown in Section 2.2, Puyuma RTO imposes no case restriction on the YP, nor does it exhibit sensitivity to islands. At the same time, however, it must satisfy the aboutness condition, i.e., the content of the CP must be “about” the referent of the XP, a constraint that is commonly observed between base-generated left-dislocated phrases and the clauses that follow them. As can be seen in the following example, the failure to satisfy the aboutness condition results in unacceptability in Puyuma RTO (15).

- (15) *ma-tiya=ku *kan Atrung* [dra trakaw dra paysu i Sawagu].
 AV-dream=1SG.ABS *SG.OBL A* [C <AV>steal ID.OBL money SG.ABS S]
 (‘I dreamt (about Atrung) that Sawagu stole money.’)

PUYUMA

Based on these observations, we argue that XPs are base-generated left-dislocated adjuncts to the CP.

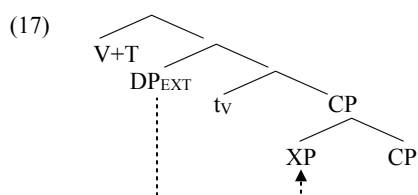
3.2. The Matrix Behavior of XPs

As discussed in Section 2.1, XPs in Formosan RTO exhibit matrix object-like behavior in terms of anaphora binding and case marking. In this subsection, we provide our accounts for these behaviors of XPs based on the proposed analysis of RTOs as embedded left-dislocations.

First, as already shown in ((4)-(6)), when XPs are reflexive pronouns, they are obligatorily bound by the matrix subjects. We argue that the reflexivization facts follow straightforwardly from the proposed structure of RTO in (14), according to which the XP adjoined to the CP is the closest DP c-commanded by the matrix subject. For concreteness, we adopt the A-chain analysis of reflexives proposed by Reinhart and Reuland (1993), which is based on the following definition of an A-chain.

- (16) *An A-chain, under its broadest definition, is any sequence of coindexation that is headed by an A-position and satisfies antecedent government; that is, each coindexed link, except for the head, is c-commanded (i.e., m-commanded) by another link, and there is no barrier between any two of the links. (693).*

According to (16), if an XP is pronominal, it must be reflexivized, as it is the closest DP that the subject DP, i.e., the head of the A-chain, c-commands.



Second, as all the examples of RTOs in this paper have shown, the XP bears morphological case that would be borne by a direct object. We argue that such case marking of XPs follows from two independently motivated generalizations of case in Formosan languages. First, CPs are “case-licensed” in Formosan languages, *i.e.*, *CPs enter into an Agree relation with a case-licensing head in these languages*. Second, as mentioned earlier, in Formosan languages, *there must be one and only one ABS-marked phrase in a given clause*. As will be discussed below, these two generalizations lead us to propose that the morphological case borne by XPs is a reflex of the case assigned to the co-occurring CP, *i.e.*, the XP inherits the case that is assigned to the *XP-CP* constituent.

Empirical motivations for the first generalization come from the A'-extraction asymmetry that has been reported in several Philippine-type Austronesian languages. According to Chung (1991, 1994, 1998), in Chamorro, the extraction of a *wh*-phrase is accompanied by verbal agreement (“*wh*-agreement”) on the verb, which indicates the case status of the *wh*-phrase. In (18), the embedded verb has an infix *<in>*, which indicates that the *wh*-phrase extracted bears OBL case. In the same example, the matrix verb has a prefix *um-*, which corresponds to nominative case. That the only element that could bear nominative case in (18) is the CP motivates the argument that Chamorro verbs agree with CPs in terms of case.

- (18) *Hayi* um-istotba si-Juan [*eci* ni m<in>ahalang i asagua-ña *eci*]? (CHAMORRO)
who; NOM-disturb PN-J [*eci* C OBL.lonely the spouse-AGR *eci*]
 ‘Who does it disturb Juan that his wife is lonely for?’ (Chung 1994:14)

A similar proposal has been put forward for Tagalog (e.g., Rackowski and Richards 2005; Law 2014), as exemplified in the following Tagalog relative clause data:

- (19) a. Gusto ko ang *libro*-ng [s<in>abi ni Fred [na b<in>ili ni Maria *eci*]].
 like 1SG ABS *book*-LK [<PV.PERF>say ERG F [LK <PV.PERF>buyERG M *eci*]]
 ‘I like the book that Fred said that Maria bought.’
 b. *Gusto ko ang *libro*-ng [nag-sabi si Fred [na b<in>ili ni Maria *eci*]].
 like 1SG ABS *book*-LK [AV.PERF-say ABS F [LK <PV.PERF>buyERG M *eci*]]
 (‘I like the book that Fred said that Maria bought.’) (Law 2014:4; glosses ours) (TAGALOG)

In (19a-b), the head of the relative clause *libro* ‘book’ is A’-extracted from the most deeply embedded clause. As is well known, A’-extractions in Tagalog impose the absolutive-only condition, i.e., only ABS-marked phrases can participate in it. However, although *libro* ‘book’ in both (19a) and (19b) is ABS-marked inside the most deeply embedded clause, only (19a) is grammatical. According to Rackowski and Richards (2005) and Law (2014), the crucial difference between the two lies in the fact that the voice of the intermediate verb *sabi* ‘say’ is PV in (19a), but AV in (19b). Hence, the most deeply embedded CP in the sentences in (19) can be analyzed as ABS-licensed by the PV-form of the verb *sabi* in (19a) but OBL-licensed by the AV-form of the same verb in (19b). This extraction asymmetry led to the following generalization:

- (20) *Only those CPs and DPs that Agree with a phase head on independent grounds (e.g., direct objects and complement clauses) are transparent for wh-extraction.* (Rackowski and Richards 2005:582)

We argue that the same generalization applies to the three Formosan languages, although its effects are manifested slightly differently. As in Tagalog, A’-extractions in Puyuma, Amis, and Seediq obey the ABS-only constraint. Consider the following *wh*-questions from Amis.²

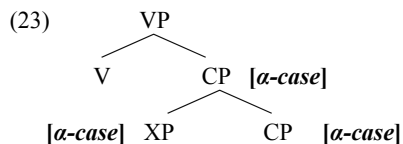
- (21) a. *cimai* ku [ka-tengil-an isu [ku fangcal-ay *eci*]]? (AMIS)
*who*_i ABS [PV.IRR-hear-LV.NMZ you.ERG [ABS pretty-AGT.NMZ *eci*]]
 ‘Who do you hear is pretty?’
 b. *cimai* ku [ka-fana’*(-an) ✓ isu/*kisu [ku fangcal-ay *eci*]]?
*who*_i ABS [AV.IRR-know-*(LV.NMZ) ✓ 2SG.ERG/*ABS [ABS pretty-AGT.NMZ *eci*]]
 ‘Who do you know is pretty?’

In (21a), the intermediate CP has the verb *tengil* ‘hear’ in non-AV form, which presumably makes the most deeply embedded CP ABS-licensed, similar to the Tagalog example (19a). As expected, the sentence is grammatical. (21b), on the other hand, involves the verb *fana* ‘know’ as the verb of the intermediate CP. Under normal circumstances, the verb *fana*’ is only compatible with AV structure, as in (22).

- (22) a. ✓ ma-fana’/fana’*-en/*-an kisu/*isu [Ø ma-ulah kaku ci Mayaw-an]. (AMIS)
 ✓ AV-know/know*-PV/*-LV 2SG.ABS/*2SG.ERG [C AV-like 1SG.ABS PN M-OBL]
 ‘You know that I love Mayaw.’

However, despite the fact that the verb *fana*’ ‘know’ normally takes an AV marker (22), in A’-extraction sentences like (21b), the same verb is observed to obligatorily co-occur with a non-AV nominalizer *-an*, indicating that the verb is functioning as a non-AV verb under extraction circumstances. We believe that this exceptional behavior of the verb under A’-extractions is due to the generalization in (20). Namely, if the verb behaved as it normally does, the most deeply embedded CP would be OBL-marked, hence an A’-extraction out of this CP would be impossible. Only if the verb in (21b) is non-AV-marked would the most deeply embedded CP receive ABS case and allow A’-extractions.³ This, in turn, supports the claim that CPs are “case-licensed” in Formosan languages.

Going from this generalization, we propose that (i) in Formosan RTO, XPs bear the case assigned to the co-occurring CP complement, and (ii) the XP and CP must form a constituent. Following these two proposals, we argue that the case-marking on the XP is a realization of the case assigned to the XP-CP constituent, as schematized in (23).



Supporting evidence for proposal (ii) come from the following Puyuma data (24), where XPs are adjoined to a complex DP, rather than a CP complement.

² The same pattern is observed in Puyuma and Seediq. Due to space limitations, only Amis data are presented here.

³ Importantly, the *ka*-prefix in (21a) serves as a PV marker, while that in (21b) serves as an AV marker, as evidenced by the different case statuses (ERG or ABS) of the external arguments introduced by the verbs in simple clauses.

According to Landau (2011), the configuration in (28) accounts for the cases of left-dislocation and copy-raising that require the presence of a pronoun, as well as prolepsis constructions in languages like Madurese (Davies 2005). Under this analysis, the properties of Amis RTO follow directly from the properties of topics in Amis. While an Amis topic must be co-indexed with an ABS-marked phrase, this co-reference relation does not respect syntactic locality or islands. In (29a) below, the topic *Ofad* is co-referential with the ABS-marked pronoun *cingra* ‘he’ embedded inside the adverbial clause (which can be either null or overt). As can be seen, this co-reference “skips” the closer ABS-marked phrase, *kaku* ‘I’, and also crosses an adjunct island. In (29b), the topic *Mayaw* is co-referential with the ABS-marked pronoun *cingra* ‘he’ inside the complex DP *sinpung* ‘news’.

- (29) a. *Ø-ci Ofadi_i*, tayra kaku i Busung, [anu pafli (*cingra_i*) takuwanan tu paysu].
ABS-PN Ø_i AV.go 1SG.ABS LOC B [if AV.give (*3SG.ABS*) 1SG.ABS OBL money]
 ‘Ofadi_i, I will go to Busung if he_i gives me money.’
- b. *Ø-ci Mayaw_i*, ma-tengil aku [kuna sinpung [adada (*cingra_i*)]]. AMIS
ABS-PN M_i PV-hear 1SG.ERG [ABS.that news [AV.be.sick (*3sg.abs*)]]
 ‘Mayaw_i, I heard the news that he_i is sick.’

Postulating the null topic operator inside the CP in Amis RTO, as in (30), hence accounts for both the ABS-only requirement imposed on YPs and the lack of island-sensitivity.

- (30) ma-fana’ kaku *ci Mayaw-an_i* [kuna sinpung [*Op_i* adada (*cingra_i*)]]. AMIS
 AV-know 1SG.ABS *PN M-OBL_i* [ABS.that news [*Op_i* AV.be.sick (*3sg.abs*)]]
 ‘I know (about Mayaw_i) the news that he_i is sick.’

4.3. Seediq RTO: Predicative CP with an A'-moved Null Operator in [Spec, CP]

Among the three languages’ RTO constructions, Seediq RTO is unique in that the XP-YP relation is island-sensitive ((12c) and (13c)). Like Amis RTOs, however, Seediq RTO imposes the ABS-only requirement on YPs ((9b) and (12b)). We argue that Seediq RTO exhibits these properties, as it involves a predicative CP and a null operator that A'-moves to [Spec, CP], as in (31).

- (31) $V_{\text{MATRIX}} \quad \mathbf{XP}_i \quad [_{\text{CP}} \text{Op}_i \quad t_i \quad]$

The structure in (31) offers a unified account for the ABS-only requirement and the island-sensitivity in Seediq RTO. As the construction involves an A'-movement, (i) the YP must be an ABS-marked phrase, as only an ABS-marked phrase can be A'-extracted, and (ii) the XP-YP relation cannot cross islands, as an A'-trace cannot be licensed with islands.

5. Conclusions and Implications

We have argued that RTO in the three Formosan languages are instances of embedded left-dislocation that involve a finite CP and a base-generated left-dislocated adjunct, the XP. As such, they involve no derived matrix objects or raising out of CPs. We have also argued that the micro-variation observed in RTOs across these three closely related languages receives a unified account under the hypothesis that a CP can establish a relationship with a left-dislocated phrase in three different ways: (i) via the *aboutness condition*, (ii) via a *base-generated null operator*, and (iii) via an *A'-moved null operator*. If the present analysis is on the right track, RTO in the three languages provides novel support for Landau’s (2011) claim that *a CP and a CP-external phrase may establish a connection via three different strategies*.

The present proposal adds RTOs in Puyuma, Amis, and Seediq to the growing list of languages with RTO-like phenomena that do not involve movement out of finite CPs, including Niuean and Fijian RTO (Massam 1985), Malagasy RTO (Paul and Rabaovoloana 1998; Pearson 2005), Tsez long-distance agreement (Polinsky and Potsdam 2001), Madurese prolepsis (Davies 2005), and Sundanese prolepsis (Kurniawan 2011). What is conspicuously missing in this picture of Formosan RTO is a strategy that involves actual syntactic movement out of CPs. Only a handful of languages with RTO-like phenomena have been argued to involve syntactic movement of an XP out of finite CPs (e.g., Japanese: Kuno 1976 and Tanaka 2002; Korean: Yoon 2007; Romanian: Alboiu and Hill 2013; and Zulu: Halpert and Zeller to appear). The lack of actual raising out of finite CPs in typologically and geographically diverse

languages, including the three Formosan languages examined in this study, suggests the pervasiveness of the generalization that A-movement out of finite CPs is prohibited (e.g., Tensed-S Condition; Chomsky 1973), and seems to be consistent with the intuition that *finite CPs form an independent and complete unit with respect to syntactic operations* (Phase Impenetrability Condition (Chomsky 2000, 2001, 2008)). The existence of the languages that arguably do allow A-movement out of finite CPs, on the other hand, raises the question of under what circumstances a language may allow such an operation.⁴

References

- Aissen, Judith. 1992. Topic and focus in Mayan. *Language* 68:43-80.
- Alboiu, Gabriela & Virginia Hill. 2013. The case of A-bar ECM: Evidence from Romanian. In Stefan Keine & Shayne Sloggett (eds.), *Proceedings of the Forty-Second Annual Meeting of the North East Linguistic Society*: 25-39. Amherst, MA: GSLA.
- Anagnostopoulou, Elena. 1997. Clitic left dislocation and contrastive left dislocation. In Elena Anagnostopoulou, Henk van Riemsdijk & Frans Zwats (eds.), *Materials on left dislocation*, 151-192. John Benjamins: Amsterdam.
- Blust, Robert. 1999. Subgrouping, circularity and extinction: Some issues in Austronesian comparative linguistics. In Elizabeth Zeitoun & Paul Li (eds.), *Selected papers from the Eighth International Conference on Austronesian Linguistics*, 31-94. Taipei: Institute of Linguistics, Academia Sinica.
- Bruening, Benjamin. 2001. Raising to object and proper movement. Ms., University of Delaware.
- Chomsky, Noam. 1977. On wh-movement. In Peter W. Culicover, Thomas Wasow, & Adrian Akmajian (eds.), *Formal syntax*, 71-132. New York: Academic Press.
- Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels, & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89-156, Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (eds.), Ken Hale: *A life in language*, 1-52, Cambridge, MA: MIT Press.
- Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria L. Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133-166, Cambridge, MA: MIT Press.
- Chung, Sandy. 1991. Sentential subjects and proper government in Chamorro. In Carol Geogopoulos & Roberta Ishihara (eds.), *Interdisciplinary approaches to language: Essays in honor of S.-Y. Kuroda*, 77-99. Amsterdam: Kluwer Academic.
- Chung, Sandra. 1994. Wh-agreement and “referentiality” in Chamorro. *Linguistic Inquiry* 25:1-44.
- Chung, Sandra. 1998. *The design of agreement: Evidence from Chamorro*. Chicago and London: University of Chicago Press.
- Cinque, Guglielmo. 1990. *Types of A'-dependencies*. Cambridge, MA: MIT Press.
- Davies, William. 2005. Madurese prolepsis and its implications for a typology of raising. *Language* 81, 645-665.
- Halpert, Claire & Jochen Zeller. To appear. Right dislocation and raising-to-object in Zulu. *The Linguistic Review*.
- Heycock, Caroline. 1993. Syntactic predication in Japanese. *Journal of East Asian Linguistics* 2:167-211.
- Kuno, Susumu. 1976. Subject raising. In Masayoshi Shibatani (ed.), *Syntax and semantics 5: Japanese generative grammar*, 17-50, New York: Academic Press.
- Kurniawan, Eri. 2011. Does Sundanese have prolepsis and/or raising-to-object constructions? *Proceedings of the 18th Annual Meeting of the Austronesian Formal Linguistics Society (AFLA 18)*, 16-30. Online publication hosted by the University of Western Ontario.
- Kuroda, Sige-Yuki. 1986. Movement of noun phrases in Japanese. In Takashi Imai & Mamoru Saito (eds.), *Issues in Japanese linguistics*, 229-272. Dordrecht: Foris.
- Landau, Idan. 2011. Predication vs. aboutness in copy raising. *Natural Language and Linguistic Theory* 29:779-813.
- Law, Paul. 2014. The syntax of Tagalog relative clauses. Paper presented at the 21st Annual Meeting of the Austronesian Formal Linguistics Society (AFLA 21). Honolulu, May 23-25, 2014.
- Massam, Diane. 1985. Case theory and the projection principle. Ph.D. dissertation, Massachusetts Institute of Technology.
- Paul, Ileana & Lucie Rabaovololona. 1998. Raising to object in Malagasy. In *The structure of Malagasy, vol. 2. UCLA Occasional Papers in Linguistics* 20:50-64.
- Pearson, Matt. 2005. The Malagasy subject/topic as an A'-element. *Natural Language and Linguistic Theory* 23(2), 381-457.
- Polinsky, Maria & Eric Potsdam. 2001. Long-distance agreement and topic in Tsez. *Natural Language and Linguistic Theory* 19, 583-646.
- Rackowski, Andrea & Norvin Richards. 2005. Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry* 36:565-599.
- Reinhart, Tanya & Eric Reuland. 1993. Reflexivity. *Linguistics Inquiry* 24:657-720.
- Saito, Mamoru. 1982. Case marking in Japanese: A preliminary study. Ms., MIT, Cambridge, MA.
- Tanaka, Hidekazu. 2002. Raising to object out of CP. *Linguistic Inquiry* 33:637-652.
- Yoon, James. 2007. Raising of major arguments in Korean (and Japanese). *Natural Language and Linguistic Theory* 25(3):615-653.

⁴ According to Bruening (2001), Passamaquoddy RTO may be added to this list, as it arguably utilizes two different strategies in establishing the XP-YP relation, one of which involves an XP base-generated at the left edge of the CP, and undergoing A-movement to the matrix domain.