Raising to Object Out of CP as Embedded Left Dislocations: Evidence from three Formosan languages

INTRODUCTION: Three closely related Formosan languages, Puyuma, Amis and Seediq, all exhibit what appear to be instances of Raising-to-Object-out-of-CP (ROC) ((1)-(3)). ROCs in these languages share superficial similarities, as they all involve the “raised” phrase (the XP) thematically belonging to an embedded finite CP that optionally appears in the matrix domain. Further, in all three languages ROCs unrestrictedly apply to all knowledge/perception verbs that subcategorize for a finite CP.

(1) aparu=ku i Arasip [dra m-uka ec\_i Senden]. PUYUMA
   `I forget that Senden has been to Arasip.'

(2) ma-fa'ana' kaku tuna wacu [ø ma-palu ni wawa (cingra)]. AMIS
   `I know that the child beat that dog.'

(3) spi-an Hana ka Watan [ø s<\m>ipaq laqi=na\_i ec\_]. SEEDIQ
   `Hana dreamt that Watan beats his child.'

In this paper, we first show that, despite the similarities, ROCs in these languages impose distinct restrictions on how the XP is associated with the finite CP. We then propose that the observed microvariation is best accounted for by analyzing these apparent ROCs as cases of embedded left dislocations that involve three independently motivated strategies: (i) concatenation of the XP and a propositional CP (4a), (ii) coindexation between the XP and an operator base-generated inside a predicative CP (4b), and (iii) coindexation between the XP and an operator that undergoes A'-movement inside a predicative CP (4c) (e.g., Aissen 1992; Culicover & Jackendoff 1999; Landau 2011).

SIMILARITIES: ROCs in all three languages involve a finite embedded CP and the “raised” XP that is in the matrix domain. The finite CP analysis for the ROC complements is motivated by the unrestricted aspect marking (5) and embedded voice type ((2), (6)). The status of the XP as a matrix element is suggested by (i) word order in (5)-(6), in which the XP precedes both overt complementizers and matrix elements, (ii) unambiguous reflexive binding (6), and (iii) the Case licensing of the XP ((2), (5), (10)), which is inconsistent with the XP being an embedded element.

(5) ma-ladram kan lsaw\_i i Senden [*(dra) adri d<em>a-deru ec\_ dra bujir]. PUYUMA
   `Senden knows that Isaw is not cooking taros now.'

(6) ma-lemed ni Kul\_as, cing\_ra, inacila [ø ma-palu ni Mayaw\_k ec\_]. AMIS
   `Yesterday Kul\_as dreamt that Mayaw\_k beats him.'

HOW THEY DIFFER: First, ROCs in Amis and Seediq require that the XP be identified with an embedded absolutive phrase (the ABS-only constraint) ((7), (8)) while such restriction is absent in Puyuma ((1), (9)). In Puyuma, the XP can be identified with any type of embedded element including absolutive (5), ergative, oblique (1), and adverbial (9) as long as it is definite (e.g., (1)).

(7) *ma-tawal aku i Kalingku [ø taira ec\_ ci Lisin]. AMIS
   `I forget that Lisin has been to Kalingku.' (cf. (1))

(8) *s<\m>un=longi=ku Skangki [ø me-n-sa ec\_ ka Ikung]. SEEDIQ
   `I forget that Ikung has been to Skangki.' (cf. (1))

(9) ma-ladram=ku an miranang na bira' [*(dra) wa-ruma=yu]. PUYUMA
   `I know that you will be back when the leaves turn yellow.'
Such oblique case in Formosan languages is the analysis of Formosan languages. how they proceed (i.e. the ban on improper movement).

restricted domains in which grammatical operations operate (i.e. Phase Impenetrability Condition) and restricted set of grammatical strategies.

how “singling out” an element in a finite embedded CP is achieved in different languages using a novel evidence for the claim that CPs are not.

The differences among ROCs in these languages derive from the putative differences in the internal structure of the embedded CPs. First, the fact that the relationship between the XP and the embedded CP is unrestricted in Puyuma can be captured under the analysis that the embedded CPs in Puyuma ROCs are propositional CPs simply concatenated with the XP (4a). For Amis and Seediq ROCs, we propose that the XP is coindexed with a null operator (Op) inside a predicative CP. The ABS-only constraint on the XP comes from the fact that Op must be identified with an absolutive phrase in both languages. Under the proposed analysis, the divergence in island-sensitivity between Amis and Seediq ROCs is accounted for as follows: In Amis, Op is base-generated as an internal topic in [Spec, CP] and unselectively binds an embedded absolutive DP (4b). Importantly, topics in Amis are observed to unselectively bind any absolutive argument regardless of syntactic locality. Thus, postulating Op as a base-generated topic inside the CP accounts for the lack of locality constraint with Amis ROCs in (13) while maintaining the ABS-only constraint. The base-generated Op analysis for Amis ROCs is further supported by the fact that an overt embedded pronoun is used in ROCs in the Central dialect of Amis (2) (Wu 2000), while traces in A-abs operations are never spell-out as pronouns in the same dialect (Wu 2000). In Seediq, Op A’-moves to [Spec, CP] (4c) and hence obeys island constraints ((10), (11)). Together, the three languages demonstrate how an embedded CP and a left dislocated XP can be associated via three distinct strategies in (4).

The microvariation in ROCs among the three languages presents novel evidence for the claim that CPs are not natural predicates and Op is necessary to make them predicative (Rothstein 1991; Landau 2011). This study also adds these three Formosan languages to the growing list of languages with ROCs without true “raising” out of CP, and provides a window into how “singling out” an element in a finite embedded CP is achieved in different languages using a restricted set of grammatical strategies. In particular, they provide further support for the notions that restrict domains in which grammatical operations operate (i.e. Phase Impenetrability Condition) and how they proceed (i.e. the ban on improper movement). Last, they present important implications for the analysis of Formosan languages. In particular, the Case status of the XP in ROCs indicates that oblique case in Formosan languages is not an inherent case, but is structurally licensed (e.g. by v). Such analysis is supported by the data from restructuring phenomena in the same languages.

**REFERENCES**