



**HAL**  
open science

## Information Structure in Northern Amis: A Morphosyntactic Analysis

Isabelle Brill

► **To cite this version:**

Isabelle Brill. Information Structure in Northern Amis: A Morphosyntactic Analysis. *Oceanic Linguistics*, 2016, *Oceanic Linguistics*, 55 (2), pp.449-479. 10.1353/ol.2016.0021 . halshs-01418746

**HAL Id: halshs-01418746**

**<https://shs.hal.science/halshs-01418746v1>**

Submitted on 16 Dec 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Information Structure in Northern Amis: A Morphosyntactic Analysis

Isabelle Bril

LACITO-CNRS AND LABEX EFL, PARIS

This paper is a study of Information Structure (IS) in Northern Amis and its interaction with the voice system. These two domains are distinct, but interacting: the voice system obeys semantic and syntactic constraints, while IS and its related discourse categories (topics, foci) follows discourse parameters. A morphosyntactic analysis of the distinctive features of topics and focusing devices is conducted, showing how pragmatics maps onto syntax. Among focusing devices are evidential *in situ* foci, which recruit two evidential forms, *han(=tu)* and *sa(=tu)*, and cleft narrow foci (informative or restrictive). Evidential foci also express stance and do not trigger any syntactic restructuring, in contrast with extracted cleft foci. As in other Formosan languages, a cleft core argument must be the syntactic pivot of the clause containing the presupposition, while there is no such constraint for cleft peripheral arguments and for adjuncts, which trigger a complement clause.

**1. INTRODUCTION.**<sup>1</sup> This study of information structure (IS) in Northern Amis has been conducted on collected fieldwork data.<sup>2</sup> They are spontaneous, recorded productions (mostly stories) from male and female speakers over sixty years of age. A few short sample originating from these spontaneous speech productions have also been recorded as elicited material.

Beyond the analysis of the specific morphosyntactic strategies expressing left-dislocated topics and various types of foci, the questions addressed are concern with the way in which pragmatically marked focus constructions interact with voice and with the privileged syntactic argument (PSA) or pivot.

There is a long tradition of analysis of the voice system in Philippine-type languages in terms of informational saliency that led to the pivot argument being labeled as “focus,” declined as Actor, Undergoer, Locative, and Instrumental Focus. Even though the syntac-

---

1. This research is supported by the LACITO-CNRS and mostly financed by the research strand 3 “Typology and dynamics of linguistic systems” of the Labex EFL (Empirical Foundations of Linguistics) (ANR-10-LABX-0083/CGI). I wish to thank three anonymous reviewers for their comments and feedback. I am very grateful to Elizabeth Zeitoun for her comments and suggestions on a previous version of this article and, more generally, for helping me start research on Northern Amis, introducing me to Amis people in Hualien, and for our collaboration in the early stages of fieldwork. Lastly, I wish to express my gratitude to my Amis informants and friends for their invaluable collaboration.

2. Unless otherwise indicated, all data are from my own fieldwork, conducted in all three communities over a period of 10 months.

tic pivot is the PSA, this system is different from information structure and from pragmatically driven constructions such as left-dislocated topics, evidential foci, and cleft foci, which operate at distinct levels.

The layout of the paper is as follows. Section 2 provides some brief background information on the voice system and IS constructions, teasing apart the syntactic-semantic parameters that select the choice of voice and argument structure from the pragmatic, discourse effects of IS. Sections 3–5 detail the strategies of IS, especially the encoding of topics (frames) and foci (restrictors). Section 6 discusses the interaction of syntax, voice, and pragmatics in focus constructions, and teases apart cleft core arguments from cleft peripheral arguments and adjuncts. Section 7 concludes.

**2. THE VOICE SYSTEM AND INFORMATION STRUCTURE IN NORTHERN AMIS.** The present analysis bears on northern Amis, the variety spoken in and around Hualien by three communities living in Natauran, Pokpok, and Lidaw. Though the Amis claim approximately 150,000 people, the actual number of speakers is much lower and declines dramatically in the younger age group. Amis belongs to the putative Eastern Formosan subgroup of the Austronesian family (Blust 1999; Sagart 2004; Ross 2009); it is spoken along the eastern coast of Taiwan and subdivides into four main dialects: northern, central, coastal, and southern (Tsuchida 1988).<sup>3</sup> These dialects display significant differences in their phonology, lexicon, and morphosyntax. Previous studies mostly concerned verbal classification and the voice system of the central dialect (Wu 2006), of Fataan Amis (Tsukida 2008), and of the Natauran dialect in the north (Chen 1987), the one analyzed here.

A brief note on the phonology and spelling system is necessary here. There is no vowel length opposition, except for expressive purposes: a lengthened vowel is expressive and often focal; it is marked by a colon. The glottal stop appears phonetically before word-initial vowels and after word-final vowels; it is rule-driven, not phonemic, and is not written. In word medial position, in the context of two adjacent vowels, the glottal stop is phonemic and occurs between them, as in *litaul* [litaʔul] ‘egg’, *uyaan* [ujaʔan]; in the context where /i/ or /u/ precede another vowel, the glottal stop is also phonemic and contrasts with glide insertion, as in *tuud* [tuʔud] ‘all’ vs. *itakuwan* ‘(to) me’, *kumiut* [kumiʔut] ‘chayote shoot’ vs. *kuniyu* ‘land snail’, *liel* [liʔəl] ‘neck’ vs. *liyal* ‘sea’. The epiglottal stop /ʔ/ is written <q>. In word-final position, /ʔ/ is produced with an aspirate release [ʔʰ], like all unvoiced final stops (*p*, *t*, *k*) (Edmonson et al. 2006). In syllable- or word-final position, /h/ is produced as a pharyngeal fricative [ħ]: consider the minimal triplet *papa* [papaʔ] ‘spoon’, *papaq* [papaʔʰ] ‘chew’, *papah* [papaħ] ‘leaf’.<sup>4</sup> Stress falls on the final syllable.

**2.1 SOME GENERAL FEATURES OF THE VOICE SYSTEM.** Amis has basic predicate initial word order. Information Structure (IS) and its related discourse categories (topic, focus) trigger modifications of that basic word order: topics are left-dislocated with

3. Sakizaya, treated by Tsuchida as a fifth dialect, is now considered as a distinct language.

4. Other symbols: <d> is [ɟ]. <b> is [β]. <l> is a retroflex flap [l̠], and <ng> is [ŋ].

a pause at their rightmost boundary, while cleft foci shift constituents around within the predicate–argument template. Additional morphemes and prosodic variation are also involved.

**2.1.1 The voice system.** Like some other Formosan languages, Amis has a complex voice system. The two basic voices are traditionally labeled Actor Voice (AV) and Undergoer Voice (UV), and the affixes are *mi-* ‘AV’ and *ma-* ‘UV’. These two morphemes also distinguish two verb classes, activity *mi-* verbs from nonactivity *ma-* verbs. *Ma-* verbs also include cognition verbs (*ma-banaq* ‘know’), psych, position, motion verbs,<sup>5</sup> states and properties: all express experiencer- or patient-oriented events. The third main class is the bare  $\emptyset$ -verb class, which hosts neither of these two basic voice prefixes; it includes adjectival, stative, property-denoting verbs, and a few deictic motion verbs (*tayni* ‘come, *tayra* ‘go’).

The UV suffix *-en*, labeled passive by Chen (1987),<sup>6</sup> occurs on all verb classes (*mi-*, *ma-*, and  $\emptyset$ -) and occurs in declarative, imperative, and hortative clauses. It derives verb forms with an undergoer nominative pivot and an optionally expressed genitive agent that must be animate, as in *adada-en cira* (hurt-UV NOM.3SG)<sup>7</sup> ‘he was hurt’ (*adada* is a  $\emptyset$ -verb). In declarative sentences, *V-en* stems express the resultant state of some action or event conducted on an affected patient by an agent, as in *palal-en=aku cira* (wake-UV=GEN.1SG NOM.3SG) ‘he was woken up by me’. UV *-en* also occurs in imperative and hortative moods, as in *kilim-en!* (lit. let it be looked for) ‘look for it!’, or hortative *tangic-en=ita!* (cry-UV=GEN.1PL.INCL) (lit. let [him] be implored by us) ‘let’s implore him’. *Ma-* constructions, on the other hand, are restricted to declarative sentences; in imperative and hortative moods, *ma-* appears as *ka-* (and *mi-* as *pi-*): compare *ma-butiq cira* ‘(s)he’s asleep/sleeping’ with *ka-butiq!* ‘go to sleep!’. Another important difference between *ma-V* and *V-en* is that the former is neutral in terms of intentionality and agent animacy, while *V-en* refers to an intentional action conducted by an animate agent. For verbs accepting both *ma-V* and *V-en* constructions, the difference is semantic, and contrasts unintentional vs. intentional agent, as shown in (1a–c). Also consider (1d), which promotes the undergoer as pivot, while the agent is optional.

- (1) a. **Mi-**pesiq t-u litaul cira.  
 AV.crush OBL-NM egg NOM.3SG  
 ‘He’s crushing (the) egg(s).’

5. Motion verbs are a type of bodily function (like ‘sneeze’, ‘cough’, ‘give birth’, and so on) and are not conceived of as activities despite their dynamic feature. Natural events (flooding, earthquake, etc.) are also affixed by *ma-*.

6. Reconstructed by Wolff as Proto-Austronesian \*-en ‘patient voice’ (1973:73).

7. Abbreviations used include ABIL, abilitative; ANAPH, anaphoric; ASS, assertive; CA-, Ca-reduplication; CAUS, causative; CMP, complementizer; DX1, proximal deictic; DX2, medial deictic; DX3, distal deictic; EPIS, epistemic; EVID, evidential; EXCL, exclusive; EXS, existential; FR, free (pronoun); FUT, future; GEN, genitive; INCL, inclusive; INST.V, instrumental voice; LOC.V, locative voice; LNK, linker; MODF, modifier; NEG, negation; NFIN, nonfinite; NM, noun marker; NOM, nominative; NMZ, nominalizer; OBL, oblique; PM, personal marker; PFT, perfect; PL, plural; PROC, processual derivation; PST, past; QM, question marker; RC, relative clause; RED, reduplication; SG, singular; SUP, superlative; TOP, topic.

Example sentences from stories I collected are sourced with the name of the story and a line number, as for example (Frog story.110) in (1d). Examples without any overt source-marking are my own recorded or elicited data.

- b. **Ma**-pesiq numaku<sup>8</sup> k-u litaul.  
 UV-crush GEN.1SG NOM-NM egg  
 ‘I crushed the egg(s).’ (by accident)
- c. Pesiq-**en** numaku k-u litaul.  
 crush-UV GEN.1SG NOM-NM egg  
 ‘The egg(s) was/were crushed by me.’ (on purpose)
- d. Tuwa wacu iri, tebuy-**en**=tu (n-ira) a mi-lakec.  
 then dog TOP carry.on shoulder-UV=PFT GEN-3SG CMP AV-ford  
 ‘As for the dog, it has been carried (by him) to ford (river).’  
 (Frog story.110)

The applicative voices are the *sa*- instrumental voice with a nominative instrument pivot, and the locative voice marked by *-an*, whose nominative pivot is locative in a broad sense. These voices apply to already derived verb stems: *mi*- stems are derived as instrumental *sa-pi*- forms, *ma*- stems are derived as *sa-ka*- forms. Nominative pivots are marked by *k*-, genitive agents by *n*-, themes and oblique arguments by *t*- (Chen 1987; Huang 1995; Starosta 1995). AV *mi*- activity verbs display a nominative Actor pivot and a theme marked by *t*-. The nominative pivot of *ma*- verbs is a nonactor in a broad sense, subsuming experiencer, seat of properties, and Undergoer. In all other voices, INST *sa*-, LOC *-an*, UV *-en*, the nominative pivot is the corresponding semantic argument, while the Agent (if expressed) is in the genitive. The voice system is thus based on the coselection of a nominative pivot, the preferred syntactic argument, and a voice affix whose semantics matches the semantics of the pivot.

**2.1.2 Parameters of voice selection and their constructions.** The two basic voice markers (AV *mi*-, UV *ma*-) are selected along various parameters: semantic verb classes (Chen 1987), telicity, and to a more limited extent, definiteness and referentiality.

**Verb classes and transitivity effects.** Verb stems fall into four main classes: those that only accept *mi*- forms, those that only accept *ma*-, those that accept both with different semantics, and those that accept none of these. *Mi*-verbs are Actor-oriented activity verbs, while *ma*-verbs are nonactivity verbs whose pivot is an Undergoer or experiencer. Alignment varies with transitivity. The Actor pivot of *mi*-verbs is nominative, the theme of extended intransitive *mi*-verbs is oblique (marked by *t*-), as in (2a). The nominative pivot of intransitive *ma*-verbs is a Non-Actor (subsuming experiencer, the seat of some property or state, entities engaged in motion). Some intransitive *ma*-verbs have an extended intransitive construction (that is, psych, ingestion, cognition verbs) with a nominative experiencer-undergoer and an oblique theme, as in ‘he’s afraid of spirits’ in (3), thus behaving like extended intransitive AV *mi*-verbs such as (2a). Intransitive and extended intransitive *ma*- verbs are labeled NAV. On the other hand, transitive UV *ma*-verbs have a nominative undergoer pivot and a genitive agent (marked by *n*-), as in (2b). In this approach, *ma*- is thus considered as oriented towards a nonactor or undergoer nominative pivot, while case assignment of the nonpivot argument varies with transitivity: with extended intransitive NAV *ma*- constructions (3), the theme is oblique; with transitive UV *ma*- constructions, the agent is genitive (2b).

8. The genitive agent is marked either by a clitic form =*aku* (1SG), =*isu* (2SG), etc., or by the full form in careful speech *numaku*, *numisu*, etc.

- (2) a. **Mi-ucur cira t-u wawa.**  
 AV-order NOM.3SG OBL-NM child  
 ‘He’s ordering the child.’
- b. **Ma-ucur k-aku n-u (w)ama.**  
 UV-order NOM-1SG GEN-NM father  
 ‘I’m ordered by father/father ordered me.’
- (3) **Ma-talaw cira t-u kawas.**  
 NAV-fear NOM.3SG OBL-NM spirit  
 ‘He’s afraid of spirits.’ (Chen 1987:272)

**Telicity.** Some verbs accept both AV *mi-* and UV *ma-*constructions. The choice correlates with telicity: *mi-*verbs are compatible with atelic or telic activities as in (4a), marked by perfect =*tu*. *Ma-*verb constructions such as (4b) are undergoer-oriented and telic, expressing the resultant state of a completed process, with a fully affected undergoer as its nominative pivot, and a genitive agent that may be left unexpressed.

- (4) a. **Mi-sangaq=tu haw t-u batikar k-isu?**  
 AV-make=PFT QM OBL-NM bicycle NOM-2SG  
 ‘Have you repaired the bicycle?’
- b. Hay, **ma-sangaq=tu k-iya batikar.**  
 yes UV-make=PFT NOM-ANAPH bicycle  
 ‘Yes, the bicycle has been repaired.’

**Patient definiteness.** Patient definiteness and referential status may also be at stake, though to a more limited extent. This is not quite as strong a constraint as in some Philippine-type languages. In Tagalog, for example, voice selection correlates with the patient’s definiteness: an indefinite patient/ theme triggers AV with an oblique theme and bars UV constructions.

In Northern Amis, the oblique theme of an activity AV *mi-*verb (or an <*em*>-verb) may be definite as in (5) or indefinite as in (6).<sup>9</sup>

- (5) **Mi-tadihang k-uhni itakuwan.**  
 AV-call NOM-3PL OBL.1SG  
 ‘They called me.’
- (6) **K<em>aen t-u tamdaw.**  
 <NAV>eat OBL-NM people  
 ‘(It) ate people.’ (quner/snake.0007)

Similarly, undergoer/experiencer-oriented *ma-*verbs, like psych verbs (without *mi-*forms) occurring in extended intransitive constructions, have an oblique theme that may be indefinite (7a) or definite as in (7b).

- (7) a. **Ma-talaw cira t-u kawas.**  
 NAV-fear NOM.3SG OBL-NM spirit  
 ‘He’s afraid of spirits.’ (Chen 1987:272)

9. Bare nouns denote indefinite entities, or definite entities whose reference is unique, identifiable, known, and shared. Definiteness is marked by deictic and anaphoric determiners inflected for case: *k-* NOM; *n-* GEN; *t-* OBL.

- b. **Ma-talaw k-aku ci baki-an numaku.**  
 NAV-fear NOM.1SG PM uncle-OBL GEN.1SG  
 ‘I’m afraid of my uncle.’

Transitive UV *ma*-constructions also allow indefinite nominative undergoer pivots, like the free choice pronoun *maamaanay* ‘anything’ (8a), and indefinite genitive agents, like the ontological noun *taw* ‘stranger’ (8b):

- (8) a. **Ma-melaw haw numisu k-u maamaan-ay hakiya?**  
 UV-see QM GEN.2sg NOM-NM RED.what-NMZ perhaps  
 ‘Have you seen anything?’ (lit. ‘Has anything been seen by you?’)
- b. **Ma-melaw k-aku n-u taw.**  
 UV-see NOM-1SG GEN-NM stranger  
 ‘I was seen by someone’ (or) ‘Someone saw me.’

**Avoidance of indefinite pronouns as nominative pivots.** Despite cases such as (8a,b) above, there is still a clear tendency in spontaneous speech to avoid indefinite pronouns like ‘someone, something’ as nominative pivots and to use existential constructions instead. In (9) and (10), a deverbal noun derived by *-ay* stands as the nominative argument of the existential verb *ira* (9) or of the negative existential verb *awaay* (10); in (10), the deverbal noun *tayni-ay*<sup>10</sup> modifies an indefinite ontological noun *tamdaw*.

- (9) **Ira k-u tayni-ay.**  
 EXS NOM-NM come-NMZ  
 ‘Someone is coming.’ (lit. ‘There’s a comer.’)
- (10) **Awaay hen [k-u [tayni-ay n-u/a]<sup>11</sup> tamdaw].**  
 NEG.EXS still NOM-NM come-NMZ GEN-NM/LNK person  
 ‘Nobody has arrived yet.’ (lit. ‘There is no coming person.’)

In Northern Amis, indefinite entities may thus be nominative pivots, whereas this is excluded in Tagalog where existential constructions are obligatory.

## 2.2 VOICE AND PRAGMATIC INFORMATION STRUCTURE (IS).

The voice system is distinct from pragmatic IS, which is encoded by different morpho-syntactic strategies and distinct prosodic features. A quick illustration of this is provided by (11a,b) below. (11a) is a declarative clause, with a sentence-initial verb in Locative voice followed by the genitive agent (*n-uhni*) and by its syntactic nominative pivot (*k-uyaan mawacayay*). The locative voice signals a superficially affected undergoer, that is, the target of mockeries. In (11b), the syntactic pivot is left-dislocated as a pragmatic topic, its determiner *uyaan* is then unmarked for case, and it is delimited at its right edge by the topic marker *iri*.

- (11) a. **Tungangan-an n-uhni [k-uyaan ma-wacay-ay].**  
 mock-LOC.V GEN-3PL NOM-ANAPH NAV-naked-NMZ  
 ‘Those (who were) naked/the naked ones were mocked by them.’  
 (Buduy nu Pangcah.069/Pangcah’s clothes)

10. *Tayni* is a compound deictic  $\emptyset$ -verb (accepting neither *mi-* nor *ma-* forms): *ta-* go/come’ + *ini* ‘here’.

11. Both constructions are allowed: either the genitive marker *n-u* or the linker *a*.

- b. [Uyaan ma-wacay-ay **iri**,] tungangan-**an** n-uhni.  
 ANAPH NAV-naked-NMZ TOP mock-LOC.V GEN-3PL  
 ‘As for those (who were) naked/the naked ones, they were mocked  
 by them.’

### 3. INFORMATION STRUCTURE AND INFORMATION HIERARCHY.

I now analyze information structure, and the various types of topic and focus constructions, defined as follows.

- (i) Topics are *frames* about which something, constituting the comment, is predicated. Topics are definite, referential entities, already present in discourse or in common knowledge (Lambrecht 1994).
- (ii) The focus is “the pragmatically non-recoverable element in an utterance (i.e. the assertion)” (Lambrecht 1994:207). Cleft focus constructions comprise an asserted restrictor (the focus, which saturates a variable) and the presuppositional content (Krifka 2007).

**3.1 SOME DISTINCTIVE FEATURES OF LEFT-DISLOCATED TOPICS AND CLEFT FOCI.** In Northern Amis, frames/topics and restrictors/foci display distinctive morphosyntactic and prosodic features. Leaving aside for now the details of the various types of topics and foci, here is a brief and contrastive presentation of left-dislocated topics and cleft foci.

A left-dislocated topic/frame may display a slightly rising intonation and is marked by a pause, while a cleft restrictor/focus and the presupposition constitute one prosodic unit, without pause. Topics are not in the scope of illocutionary force operators, they are neither questioned nor negated, by contrast with cleft foci.

**3.1.1 Left-dislocated topics.** Left-dislocated topics stand in the left periphery of the sentence and are unmarked for case; they are followed by a pause (marked //) and an optional topic marker *i(ri)* bounding their right edge, as in (12). There may be an intonational rise on *iri*.

- (12) Ira tataang-ay a quner **iri**//, k<em>an t-u tamdaw sa.  
 DX3 big-MODF LNK snake TOP <NAV>eat OBL-NM people say/EVID  
 ‘As for that big snake//, it ate people.’ (quner/snake.0007)

The topic in (13) is a stage-setting time frame, followed by an existential clause introducing a new indefinite entity or a new topic.

- (13) I ayaá:yaw<sup>12</sup>//, na ira k-u tataang-ay a quner,  
 LOC RED.before PST EXS NOM-NM big-MODF LNK snake  
 ‘A loo ng time ago/once upon a time//, there was a big snake,  
 k<em>an t-u tamdaw k-ina quner.  
 <NAV>eat OBL-NM people NOM-DX2 snake  
 ‘this snake ate people.’ (quner/snake.0002)

12. In reduplicated constructions, with emphatic prosody and superlative meaning, as in *ayaá:yaw* (3-mora reduplication of *ayaw* ‘before’), stress falls on the lengthened initial vowel of the root.



**3.1.2 Cleft focus.** While arguments are postverbal in a declarative sentence, as in (14a), a cleft argument stands in clause-initial position, is predicative, and has a gerund relative clause (RC) as its argument (14b,c).

- (14) a. K<em>an t-u tamdaw k-ina tumay.  
 <NAV>eat OBL-NM people NOM-DX2 bear  
 ‘This bear ate people.’
- b. [U tumay **haw**] [k-u k<em>an-ay t-u tamdaw]?  
 NM bear QM NOM-NM <NAV>eat-NMZ OBL-NM people  
 < predicate focus > < — argument containing presupposition — >  
 ‘Was it a bear that ate people?’ (lit. ‘the eater of people/the one who ate people’)
- c. Caay! [caay ka u tumay] [k-u k<em>an-ay],  
 no! NEG NFIN NM bear NOM-NM <NAV>eat-NMZ  
 [u tataang-aya quner] [k-ira k<em>an-ay t-u tamdaw sa].  
 NM big-MODF LNK snake NOM-DX3 <NAV>eat-NMZ OBL-NM people say/EVID  
 ‘No, it was not a bear that ate (them)/the eater, it’s a big snake the one that ate people.’

Cleft NPs are nonverbal predicates whose nominative argument is the nonfinite RC containing the presupposition (see Shibatani 2009 for a similar analysis). A cleft focus can be questioned and negated: like all predicates, *tumay* in (14c) is negated by *caay*, triggering the nonfinite *ka* form of nonverbal predicates and nonactivity ( $\emptyset$ , *ma-*, <em>) verb stems.

Cleft core arguments have their semantic role indexed on the verbal voice of the RC: the Actors *u tumay* (14b) and *u tataang-aya quner* (14c) are indexed by *k<em>aen* ‘eat’.

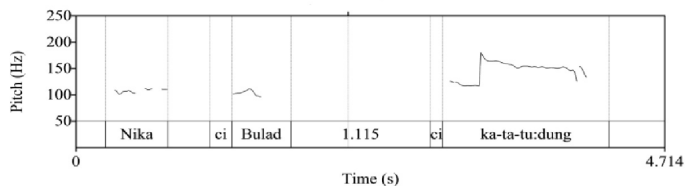
### 3.1.3 Distinctive morphosyntactic features of left-dislocated topics and foci.

The distinctive features of left-dislocated topics and cleft foci are further illustrated in (15) and (16). *Aku* in (15) is a left-dislocated topic pronoun, followed by a pause; while in (16), *aku* is focused, predicative, and followed by its nominative argument (*ku kaka*) without a pause (compare with the declarative sentence: *ci kaka=isu kaku* ‘I’m your elder sibling’).

- (15) “**Aku**//, maka-tengil (k-isu) ma-rarum=tu k-aku.”  
 FR.1SG ABIL-hear NOM-2SG NAV-sad=PFT NOM-1SG  
 ‘As for me, you can hear that I am sad.’ (Cabay=aku/my friend.00153)
- (16) “**Aku k-u kaka**, manay ci kaka=isu” han  
 FR.1SG NOM-NM older.sibling so PM older.sibling=GEN.2SG do.so  
 n-ina=aku.  
 GEN-mother= GEN.1SG  
 ‘I’m the elder one/it’s me the elder, so (I’m) your elder sibling’ my  
 mother said.’ (Cabay=aku.00144)

The following examples contrast two protagonists, Moon and Sun. In (17), *ci Bulad* is a left-dislocated topic (followed by a one second pause).

- (17) Nika ci Bulad, // caay ka-ta-tu:dung.<sup>13</sup>  
 but PM Moon NEG NFIN-CA.RED-apt  
 ‘But as for Moon, he is really unsuitable (for this).’  
 (U teker ni Adek/Adek’s bow.010)



By contrast, *ci Cidal* in (18b) is a restrictive cleft focus without pause; compare with the declarative statement in (18a).

- (18) a. Mi-kutay=tu ci Cidal.  
 AV-replace=PFT PM Sun  
 ‘Sun replaced (him).’  
 b. Ci Cidal=tu k-u mi-kutay-ay.  
 PM Sun=PFT NOM-NM AV-replace-NMZ  
 ‘It was Sun who was the substitute/replaced (him).’, or ‘Sun was the substitute.’  
 (U teker ni Adek.018)

### 3.2 DISCOURSE STRATEGIES AND THEIR COMBINATIONS.

Example (19) displays a sequence of discourse strategies: *wawa* is a detached topic (in bold), followed by a topic frame clause (in italics), both of which are continued topics with no contrastive effect; then a restrictive cleft focus *uraan a cacay* (underlined) follows, with a predicative demonstrative *uraan* is predicative, marked by the perfect marker =*tu* and a nominative nonfinite clausal argument (*k-u sa-simsim-en n-u niyam*). The undergoer role of the cleft is indexed on the passive UV *-en* verb form of the RC.

- (19) Hay **wawa** //, *uru u caay-ay hen pi-kikung,*  
 yes child that.one NM NEG-NMZ still NFIN.AV-marry  
 ‘Yes (among) the **children**, *the still unmarried one,*  
uraan=tu a cacay [k-u sa-simsim-en n-u niyam].  
 DX3=PFT LNK one NOM-NM CA RED-think-UV GEN-NM GEN.1PL EXC  
 ‘it is that one that we have concerns about.’ (lit. ‘It is that one that is our concern.’)  
 (Urip nu Balah.083-084)

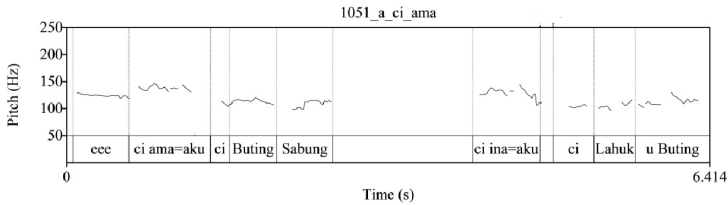
4. TOPIC CONSTRUCTIONS. Topics fall under two subtypes: (i) equative and ascriptive constructions with neutral topic-comment structure, and (ii) left-dislocated topics.

4.1 EQUATIVE AND ASCRIPTIVE CONSTRUCTIONS WITH NEUTRAL TOPIC-COMMENT STRUCTURE. Equative constructions (identity statements) such as (20) and (21) are neutral topic-comment constructions, in contrast with pragmatically marked left-dislocated topics (see 4.3). Equative constructions

13. The negation *caay* is realised as [ci], while the verb *tudung* is expressively lengthened as [tu:duŋ].

involve topic nominals and a nonverbal predicate (the comment); each segment is flagged by a noun marker (*u* for common nouns, *ci* for proper nouns).<sup>14</sup> A very short pause (marked /) occurs between the topic and comment.

- (20) **Ci** ama=aku/ **ci** Buting Sabung// **ci** ina=aku/  
 PM father=GEN.1SG PM Buting Sabung PM mother=GEN.1SG  
**ci** Lahuk u Buting.  
 PM Lahuk NM Buting  
 ‘My father was Buting Sabung, my mother was Lahuk u Buting.’  
 (Urip nu Balah.005)



Ascriptive constructions predicating a property pattern similarly.

- (21) **U** kaka/ **u** babahiy-an//, **u** saba/ **u** babainay-an.  
 NM elder.sibling NM girl-LOC NM younger.sibling NM boy-LOC  
 ‘The elder child was female, the younger was male.’ (Flooding.0011)

The topic is optionally marked by *iri*, and the construction is reversible as in (22a,b). Compare with the declarative sentence in (22c).

- (22) a. **U** saka-ulah-an numaku a radiw **iri**, **u**  
 NM SUP-like-LOC GEN.1SG LNK song TOP NM  
 ka-la-likid-an a radiw.  
 NFIN-CA.RED-dance-LOC LNK song  
 ‘As for my favorite music, it is the music of the time of dance festival.’  
 b. **U** ka-la-likid-an a radiw **iri**, **u** saka-ulah-an  
 NM NFIN-CA.RED-dance-LOC LNK song TOP NM SUP-like-LOC  
 numaku a radiw.  
 GEN.1SG LNK song  
 ‘As for the music of the time of the dance festival, it is my favorite music.’  
 c. Saka-ulah-an=aku **k**-isu.  
 SUP-like-LOC=GEN.1SG NOM-2SG  
 ‘You are my favorite.’

The equative construction in (23a) displays nominals in topic-comment relation. The comment *u pi-sadipit-an* (a deverbal noun derived from *mi-sadipit* ‘go shellfishing’) is negated by *caay* in (23b), triggering the nonfinite *ka* form also occurring with nonverbal predicates. Topics are out of the scope of illocutionary markers. Compare with a cleft focus in the scope of the negation marker in (23c).

14. The personal “article” is case-marked as absolutive/nominative *c-i* and genitive *n-i*.

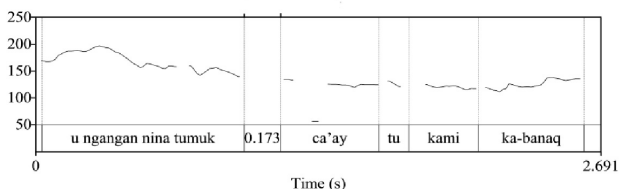
- (23) a. U sa-wali-an, u pi-sadipit-an n-u kayuin.  
 NM SA-east-LOC NM NFIN.AV-shellfish-LOC GEN-NM young.woman  
 ‘Eastward is the shellfishing place of unmarried young women.’
- b. U sa-wali-an, caay ka u pi-sadipit-an n-u kayuin.  
 NM SA-east-LOC NEG NFIN NM NFIN.AV-shellfish-LOC GEN-NM young.woman  
 ‘Eastward is not the shellfishing place of unmarried young women.’
- c. Caay ka u sa-wali-an k-u pi-sadipit-an  
 NEG NFIN NM SA-east-LOC NM-NOM NFIN.AV-shellfish-LOC  
 n-u kayuin.  
 GEN-NM young.woman  
 ‘It’s not eastward the shellfishing place of unmarried young women.’

**4.2 AUTONOMOUS, DETACHED TOPICS.** Not all detached topics are syntactic constituents (argument or adjunct) of the clause, some are “syntactically autonomous, extraclassical elements” (Lambrecht 1994: 193) standing in loose relationship with the rest of the sentence and providing a general frame. In (24), ‘forest’ is such a syntactically autonomous, loose stage frame.

- (24) [Kilakilang-an], ira k-u buhang n-iyā kilang, kay itini=tu.  
 RED.forest-OBL EXS NOM-NM hole GEN-ANAPH tree perhaps here=PFT  
 ‘In the forest, there was a hole in that tree, perhaps it (the frog) was there.’  
 (Frog story.077)

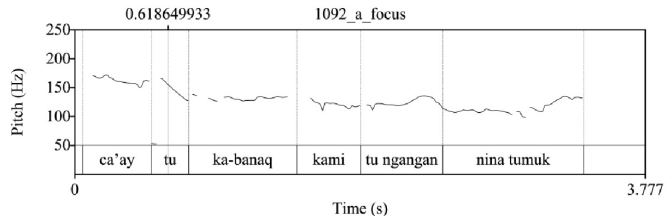
Example (25a) also displays a long, loosely connected topic, *u ngangan n-ina tumuk*. Loosely connected topics such as (25a) do not display rising intonation. Compare (25a) with the declarative sentence (25b) and its graph, the main distinctive feature is the pause.

- (25) a. U ngangan n-ina tumuk//, caay=tu k-ami<sup>15</sup> ka-banaq  
 NM name GEN-DX2 chief NEG=PFT NOM-1PL.EXCL NFIN-know  
 t-ina tumuk(-an) n-ina Cikasuan.  
 OBL-DX2 chief(-OBL) GEN-DX2 Cikasuan  
 ‘As for the chief’s name, we no longer know the chief of the Cikasuan.’  
 (Cikasuan.020)



- b. Caay=tu ka-banaq k-ami t-u ngangan n-ina tumuk.  
 NEG=PFT NFIN-know NOM-1PL.EXCL OBL-NM name GEN-DX2 chief  
 ‘We no longer know the name of that chief.’

15. *Kami* is raised; compare with the neutral order *caay=tu ka-banaq k-ami* in (25b).



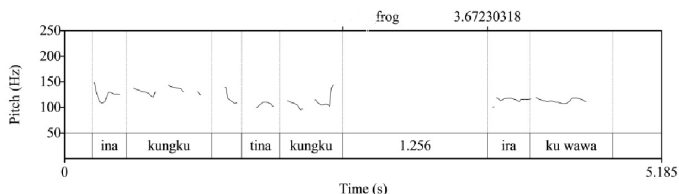
**4.3 LEFT-DISLOCATED TOPICS.** Left-dislocated topics are in sentence-initial position, followed by a pause, and optionally marked at their right edge by *i(ri)*, which is also a sequential connector ‘and then’. All arguments and adjuncts (location, time, and so on) can be left-dislocated.

The topic marker *i(ri)* is optional, as in (26), where each of the two sentences has a left-dislocated topic (*u babainay* ‘boy’ and *u babahi* ‘girl’) followed by a comment, itself structured around a focused nominal predicate (*ci Bulad* and *ci Cidal*) and a nominalised verb form marked as the nominative pivot and containing the presupposition.

- (26) [U babainay]//, [ci Bulad [k-u ni-pa-ngangan]];  
 NM boy PM Moon NOM-NM PFT NMZ-PROC-name  
 < — topic — > < — cleft focus – presupposition — >  
 ‘As for the boy, *Moon* is the name given (to him);  
 [u babahi]//, [ci Cidal [k-u ni-pa-ngangan]].  
 NM girl PM Sun NOM-NM PFT NMZ-PROC-name  
 < — topic — > < — cleft focus – presupposition — >  
 ‘the girl, *Sun* is the name given (to her).’ (U teker ni Adek.004)

**4.4 EMPHATIC LEFT-DISLOCATED TOPICS.** A particular stylistic device, illustrated in (27) and (28), emphasizes a left-dislocated topic: the NP is repeated, once in its unmarked case form (*iya tatakulaq, ina kungku*), the second time in the oblique form (*t-iya tatakulaq, t-ina kungku*), followed by a pause. An optional evidential, focal marker *sa=tu* or *han=tu*, may occur between them as in (27) (see §5.1 for more details). This is a common device in story telling. The caseless form has slightly higher pitch than the oblique form *tina kungku*, followed by a long pause.

- (27) **Iya** tatakulaq **han=tu t-iya** tatakulaq iri//,  
 ANAPH frog do.so=PFT OBL-ANAPH frog TOP  
 pa-tayra-(e)n n-**iya** wawa i puduk.  
 CAUS-go-UV GEN-DX3 child LOC bottle  
 ‘As for that *frog*, it was put by the child in a bottle.’ (Frog story.008)
- (28) **Ina** kungku **t-ina** kungku//, ira k-u wawa, ira k-u wacu,  
 DX2 story OBL-DX2 story EXS NOM-NM child EXS NOM-NM dog  
 ira k-u tatakulaq.  
 EXS NOM-NM frog  
 ‘As for the *topic* of this story, there is a child, a dog and a frog.’  
 (Frog story.003)



**4.5 THE SYNTAX OF LEFT-DISLOCATED TOPICS.** Left-dislocated topics are unmarked for case. There may be resumptive traces of their function in the clause; but unlike cleft foci, there is no constraint on their being syntactic pivot indexed by voice in the comment. In (29a), the left-dislocated NP is coreferential with the 3rd singular possessive adjective *n-ira*; compare this with the neutral declarative sentence (29b):

- (29) a. Ina kapah ci Adek/, ira haw k-u balucuq n-ira.  
 DX2 youth PM Adek EXS EPIS NOM-NM heart GEN-3SG  
 ‘As for that young Adek, he had a plan.’ (U teker ni Adek.057)
- b. Ira haw k-u balucuq n-ina kapah ci Adek.  
 EXS EPIS NOM-NM heart GEN-DX2 youth PM Adek  
 ‘That young Adek had a plan.’

In (30a), the left-dislocated agent is unmarked for case (*uhni* is a caseless free pronoun), while in the declarative sentence (30b) the agent *n-uhni si-kawas-ay* is genitive.

- (30) a. **Uhni** si-kawas-ay/, ma-araw=tu k-iy a isaw.  
 FR.3PL get-spirit-NMZ UV-see=PFT NOM-ANAPH surely  
 ‘They, the shamans, (they) can see that surely.’ (lit. ‘As for shamans, this is seen [by them] surely.’) (Cabay=aku.00116)
- b. Ma-araw=tu k-iy a **n-uhni** si-kawas-ay isaw.  
 UV-see=PFT NOM-ANAPH GEN-3PL get-spirit-NMZ surely  
 ‘The shamans can surely see that.’ (lit. ‘That is seen by them the shamans).’

Topic adjuncts can be left-dislocated with their prepositions:

- (31) ... **i** putal **iri**, ta, ma-peciq k-iy a kureng.  
 LOC outside TOP then UV-break NOM-ANAPH jar  
 ‘... outside, then the jar was broken.’ (Frog story.047)

**4.6 CLAUSAL TOPICS WITH *iri*: BACKGROUNDING AND FOLLOW-UP EVENTS.** Left-detached clausal topics, optionally marked by *i(ri)*, are pragmatically backgrounded. They are frames for the clause containing the assertion and the follow-up events. Time or causal frame clauses are adjuncts of the main clause, in a relation of “ad-clausal subordination” (van Valin 2005:75). Prosodic analyses show a pause after the topic clause, with often flat, sometimes decreasing intonation indicating follow-up events.

- (32) [Sulinay **mi-kilim** k-iy a wacu **iri**/], mukmuk-**an** n-ira  
 indeed AV-look.for NOM-ANAPH dog TOP stuck-LOC GEN-3SG

k-*iya*      *kureng*.

NOM-ANAPH jar

‘Indeed, (as) the dog is searching, he gets stuck into the glass jar.’ (lit. ‘the jar was his trap’) (Frog story.040)

The pragmatic, discourse effects of topic-comment constructions are distinct from the semantics and syntactic functions of voice and nominative pivot selection. The shift of pivot and voice in (32) signals a shift of saliency and perspective: AV *mi-kilim* ‘look for’ takes the Actor *wacu* ‘dog’ as its nominative pivot, while the locative form *mukmuk-an* ‘stuck’ takes *kureng* ‘jar’ as its nominative pivot. The change of voice and pivot shifts attention from an Actor engaged in some activity to a location (the jar that traps its head); the shift of syntactic pivot has a pragmatic impact, but differently from marked IS constructions.

**5. IN SITU EVIDENTIAL FOCI: THE FUNCTIONS OF QUOTATIVE SA AND SIMILATIVE han.** *In situ* evidential foci are marked by the quotative verb *sa* ‘say’ or the similative verb *han* ‘be/do thus’, which are recruited for pragmatic functions.<sup>16</sup> Both *sa* and *han* have quotative<sup>17</sup> functions, as shown by (33) and (34): their distinct feature is voice—*sa* only accepts Actor pivot, *han* accepts Actor or Undergoer pivot.

- (33) “Kaph=tu haw k-isu?” **han**=aku k-u cabay.  
 good=PFT QM NOM-2SG do.so=GEN.1SG NOM-NM friend  
 “‘Are you fine?’ I said to my friend.’ (lit. ‘my friend was told by me.’)  
 (Cabay aku.00082)

- (34) “Ta haw” **sa**=tu k-*iya* niyaruq.  
 go QM say=PFT NOM-ANAPH villager  
 “‘Shall we go?’ the villagers said.’

**5.1 EVIDENTIAL IN SITU FOCAL MARKERS.** *Sa* and *han* also serve as evidential markers signaling second-hand, hearsay, “alleged” information, sometimes also signaling personal judgment or testimony. As evidential markers, they may occur in their bare form, *sa* and *han*, as in (35)–(38); while in narratives, they mostly occur in their perfect form, *sa=tu* and *han=tu*. Both *sa* and *han* are also recruited as pragmatic, in situ focal markers. Their function, as quotative, evidential, or focal markers, is context-dependent. When used as focus markers, no syntactic restructuring is involved; focal entities do not move, they are simply under the scope of postposed *han(=tu)* or *sa(=tu)*, without pause. In contrast with cleft foci, but like topics, evidential foci with *satu*, *hantu* are neither questioned nor negated.

**5.1.1 Evidential han.** Example (35) displays an informative focus *ci Bulad*, followed by evidential *han*:

16. To my knowledge, this is the first time *sa* and *han* have been analyzed as evidential and pragmatic markers.

17. Just like *faire* ‘do’ in colloquial French (*alors, il fait “ne t’inquiète pas”* ‘so, he goes “don’t worry”’) or like the similative expression *be like* in colloquial English (*she’s like “you’re so cute!”*).

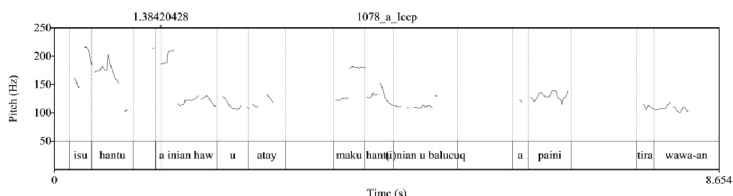
- (35) Ci Bulad **han** k-u ni-pa-ngangan n-u ma-tuas-ay  
 PM Moon do.thus NOM-NM PFT.NMZ-PROC-name GEN-NM NAV-old-NMZ  
 t-u babainay-an a wawa.  
 OBL-NM boy-LOC LNK child  
 ‘Moon, as it were, is the name given by the old people to the boy.’

**5.1.2 Focal *han*.** These verbs developing as evidential forms refer to alleged information and come to play a role as focus markers and assertion strengthening devices as in (36). This is in situ focus; *han* has focal scope on the locative phrase *i terung*, which is split from its genitive determiner.

- (36) I terung **han**<sup>18</sup> n-u pitu-ay a ta-tuki-an.  
 LOC middle do.thus GEN-NM seven-MODF LNK CA.RED-clock-LOC  
 ‘At seven o'clock sharp.’ (lit. ‘Right in the middle of seven.’)  
 (Chen 1987:109)

**5.1.3 Contrastive foci with *han=tu*.** Focal *han(=tu)* also has assertive, contrastive force, as in (37), contrasting two free pronouns, *isu* ‘you’ and *maku* ‘me’.

- (37) Isu **han=tu/** a inian haw u atay, maku **han=tu** inian u balucuq  
 FR.2SG do.thus=PFT FUT DX1 EPIS NM liver FR.1SG do.thus=PFT DX NM heart  
 < focus domain > < focus domain >  
 a pa-ini t-ira wawa-an.  
 CMP CAUS-DX1 OBL-DX3 youth-OBL  
 ‘You, (it)’ll (be) the liver, me, it’s the heart that (we’ll) present to the young girl.’  
 (Icep/betel-nut.040)



**5.1.4 Evidential and focal *sa(=tu)*.** *Sa* may also focus the verb in situ, like the evidential focus *cakay sa* ‘what (she) bought’ in (38); in the second part, the narrow focus *u qepah* is cleft and its undergoer role is indexed by the passive UV *-en* verb form in the RC.

- (38) Caay pi-cakay cira t-u miluku, cakay **sa**,  
 NEG NFIN.AV-buy NOM.3SG OBL-NM milk buy EVID  
 u qepah k-u cakay-**en** n-ira.  
 NM wine NOM-NM buy-UV GEN-3SG  
 <cleft focus> < — presupposition — >  
 ‘She did not buy milk, what (she) bought, it’s wine that she bought.’  
 (lit. ‘that was bought by her’)

In (39a), *sa* focuses the numeral predicate *in situ*; compare with the declarative clause (39b).

- (39) a. Cacay **sa** ci Adek a mi-herek t-u demak.  
 one EVID PM Adek CMP AV-finish OBL-NM work  
 ‘Alone was Adek to finish the work.’

18. Chen translates *han* as ‘exactly’ and generally treats *sa* and *han* as adverbial markers.



- b. Cacay ci Adek a mi-herék t-u demak.  
 one PM Adek CMP AV-finish OBL-NM work  
 ‘Adek alone finished the work.’

Focal evidential *sa=tu* also occurs in ascriptive (topic-comment *u* NP, *u* NP) constructions (see 4.1), such as (40). The neutral topic *u sa-wali-an*, is made focal by *sa=tu*, and tagged as alleged information. The pragmatic function of *in situ* evidentials is local, and does not trigger the kind of syntactic reorganization that clefts do.

- (40) U sa-wali-an **sa=tu**, u pi-sadipit-an<sup>19</sup> n-u kayuin.  
 NM sa-east-LOC say/EVID=PFT NM NFIN.AV-shellfish-LOC GEN-NM young.woman  
 < – EVID. focus domain –>  
 <— topic —> <————— comment —————>  
 ‘Eastward (allegedly) is the shellfishing place of unmarried young women.’  
 (U lisin.017)

Compare with the cleft locative phrase *i sa-wali-an* in (41) triggering a nominative RC.

- (41) I sa-wali-an k-u pi-sadipit-an n-u kayuin.  
 LOC sa-east-LOC NOM-NM NFIN.AV-shellfish-LOC GEN-NM young.woman  
 ‘It’s in the east the shellfishing place of unmarried young women.’

**5.2 SUMMARY CLAUSES: BACKGROUNDED TOPICS vs. EVIDENTIAL STANCE.** Summary clauses and event anaphoric clauses are also turned into frames (42) or into evidential focal constructions (43). Topic summary clauses such as *na ma-ha en iri* ‘this being so/after that’ (42) sum up a prior state of things and prepare a shift in perspective (the boy peeking in the tree-hole wonders if that is the frog’s hiding spot).

- (42) Tuwa, [na ma-haen iri], “ira k-uyu buhang n-ya tatakulaq”  
 then PST UV-be.thus TOP EXS NOM-ANAPH hole GEN-ANAPH frog  
 saan haw; caay ka u buhang n-u tatakulaq.  
 say/EVID QM NEG NFIN NM hole GEN-NM frog  
 ‘Then, this being so, “is that the frog’s hole?” he thinks; it was not a frog’s hole.’  
 (Frog story.064)

Compare with a focal evidential summary clause with strengthened assertive force (43).

- (43) Tuwa, [na ma-haen sa=tu], ma-h(a)kulung=tu k-uhni tara  
 then, PST UV-be.thus EVID/say=PFT NAV-go.together=PFT NOM-3PL go  
 i putal.  
 LOC outside  
 ‘Then, this being so, as it were, they went together outside.’ (Frog story.055)

**6. INFORMATIVE AND RESTRICTIVE FOCUS CONSTRUCTIONS IN NORTHERN AMIS.** Not all WH-questions and their answers, nor yes-no questions, involve clefting. This is discussed in 6.1; broad informative foci are presented in 6.2, and narrow cleft constructions in 6.3.

19. *Sadipit* ‘shellfish’ is derived as an activity verb *mi-sadipit* ‘get shellfish’ and is deverbalized as a location noun *pi-sadipit-an* ‘the place where one collects shellfish’.

**6.1 QUESTIONS WITHOUT CLEFTING.** Closed, yes-no questions are marked by *haw* ~ *saw* and are not cleft.

**6.1.1 Yes-no questions with *haw* ~ *saw*.** The question words *haw*~*saw* are in post-verb/predicate position (44a). Compare with the declarative sentence (44b). Cleft WH-constructions, on the other hand, involve extraction and index the semantic role of the cleft argument on the verb in the RC, as in (44c).

- (44) a. A **tayra haw** k-isu?<sup>20</sup>  
 FUT go QM NOM-2SG  
 ‘Will you go?’
- b. Na **tayni ci** Puyar **inacila**.  
 PST come NOM PM Puyar yesterday  
 ‘Puyar arrived yesterday.’
- c. Na **cima k-u** **tayni-ay** **inacila**?  
 PST who NOM NM come-NMZ yesterday  
 ‘Who arrived yesterday?’ (lit. ‘Who’s the comer yesterday?’)

**6.1.2 WH-questions without movement in nonverbal equative and ascriptive sentences.** Not all WH-question words involve clefting. In equative and ascriptive sentences, which identify or attribute a function or a property as in (45)–(47), the WH-word occurs in the standard predicate initial position; no other position is available. In (45b), the predicate *ci Dungi* is the assertion, while the nominative argument is presupposed.

- (45) a. Cima k-u **ngangan=isu**?  
 who NOM-NM name=GEN.2SG  
 ‘What’s your name?’
- b. Ci **Dungi k-u** **ngangan=aku**.  
 PM Dungi NOM-NM name=GEN.1SG  
 ‘My name is Dungi.’

No movement or clefting is involved in (46) or (47) either: the WH-words *u maan*? ‘what?’ and *u cuwa-an*? ‘which?’ and the informative focus *u kingcal* ‘policeman’ are predicates, followed by their nominative argument.

- (46) a. U **maan k-u** **demak** **numisu**?  
 NM what NOM-NM work GEN.2SG  
 ‘What work do you do? (lit. ‘Is a what your work?’)’
- b. U **kingcal k-aku**.  
 NM policeman NOM-1SG  
 ‘I’m a policeman.’
- (47) U **cuwa-an k-u** **wama=isu**?  
 NM which-LOC NOM-NM father= GEN 2SG  
 ‘Which one is your father?’

20. *Tayni* ‘arrive’ and *tayra* ‘go there’ belong to the  $\emptyset$ -verb class (that is, without *mi-* and *ma-* voice affixes).

**6.2 BROAD INFORMATIVE VERB PHRASE OR SENTENCE FOCUS WITH *maan*.** Broad verb phrase or sentence foci, whose focus domains are actions or events, do not involve movement or clefting either. *Maan* ‘what?’ is then derived as a voice-marked verb AV *mi-maan* ‘do what?’, UV *ma-maan*, passive UV *maan-en*, instrumental *sa-pi-maan* or *sa-ka-maan*.<sup>21</sup> Consider *ma-maan* ‘what happened?’ in (48a), and the sentence focus in (48b) that contains wholly new information, without any partition between assertion and presupposition (Lambrecht 1994, 2000; van Valin 2005). Sentence (48b) is also the standard declarative construction, with a sentence-initial predicate.


- (48) a. **Ma-maan** (saw) k-isu?  
 UV-do.what ASS NOM-2SG  
 ‘What’s the matter with you?’
- b. Kaw **ma-baltu=tu** k-aku n-u karireng.  
 nearly UV-collide=PFT NOM-1SG GEN-NM car  
 ‘I nearly got hit by the car.’

Compare with the AV interrogative verb *mi-maan* in (49), and its passive UV form (50).

- (49) **Mi-maan** k-isu anini?  
 AV-do.what NOM-2SG now  
 ‘What are you doing now?’ (Kasubucan.018)
- (50) **Maan-en=tu=isu** k-iya raic?  
 do.what-UV=PFT=GEN.2SG NOM-ANAPH rope  
 ‘What have you done to this rope?’

The WH-verb marked for instrument voice as *sa-pi-maan* ‘used what for?’ in (51a) questions the goal of action.

- (51) a. **Sa-pi-maan=isu** k-unian a puut?  
 INST.V-NFIN.AV-do.what=GEN.2SG NOM-DX1 LNK knife  
 ‘What do you use this knife for?’
- b. **Sa-pi-cikcik=aku** t-ina titi-an k-uraan a hawan.  
 INST.V-NFIN.AV-cut=GEN 1SG OBL-DX2 meat-OBL NOM-DX3 LNK machete  
 ‘I use that machete to cut the meat.’ (lit. ‘That machete is used by me to cut the meat.’)

**6.3 THE SYNTAX OF NARROW FOCUS: CLEFT CORE ARGUMENTS.** In contrast with WH-question words in equative and ascriptive sentences (6.1), and in contrast with broad sentence foci (6.2), narrow focus involves clefting, with a clear partition between assertion and presupposition. Narrow focus subdivides into informative and restrictive focus. Informative foci (typically answering WH-questions) identify a variable that satisfies a predication, while restrictive foci identify a variable exhaustively, sometimes contrastively (Lambrecht 1994; Van Valin 2005). All arguments can be cleft, but with distinct syntactic behavior. Cleft core arguments must be the syntactic pivot of the RC containing the presupposition, while cleft peripheral arguments (6.4.2), cleft agents (6.5) and cleft adjuncts (6.6) head a-marked complement clauses. I analyze cleft core arguments first. 

21. The interrogative pronoun *cima* ‘who?’ has no verbal derivation.

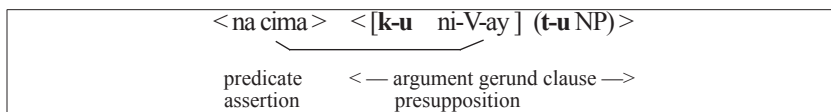
**6.3.1 Nominative only constraints on extraction and relativization.** A well-known constraint on extraction in Formosan and Philippine-type languages is that only nominative pivots can be extracted and relativized, triggering voice changes to match the semantic role of the cleft or the relativized argument (Keenan and Comrie 1977, 1979).

In (52a,b), *cima* ‘who?’ and the cleft free pronoun *cira* have their Actor role indexed by AV *mi-* on the gerund verb form (marked by *-ay*) in the RC. The RC itself is marked as substantival by a nominative determiner *k-ira* in (52a), or by the substantive morpheme *u*, case-marked as nominative *k-u* (52b). These gerund RCs are arguments of the WH-focal predicate. The gerund verb form *mi-...-ay* retains verbal properties with oblique arguments (*itisuwan*, *itakuwan*).

- (52) a. Na **cima** [**k-ira** **mi-tataq-ay** itisuwan?]  
PST who NOM-DX3 AV-cut-NMZ OBL.2SG  
 ‘Who wounded you?’ (lit. ‘Who was that wounder (of) you?’)
- b. Na **cira** [**k-u** **mi-tataq-ay** itakuwan].  
PST FR 3SG NOM-NM AV-cut-NMZ OBL.1SG  
 ‘It was him who wounded me.’ (lit. ‘It was him the wounder (of) me.’)

*Cima* must be in initial, predicate position, it may not occur *in situ* in the position of the variable (*\*Mi-tataq-ay cima itisuwan?*). The cleft focus contains the assertion and is in the scope of illocutionary markers, while the gerund nominative RC containing the presupposition is unasserted, as in figure 1.

FIGURE 1.



The cleft NP is a nonverbal predicate, possibly marked for Tense/Aspect as in (52); its nominative argument is the restrictive gerund RC. In (53a,b), the cleft NPs *u maan* and *u turun* have their undergoer role indexed on the UV *-en* form in the RC. Compare with the declarative sentence in (53c). Due to conflicting role indexing, a cleft undergoer and an AV verb are ungrammatical, as shown in (53d).

- (53) a. U maan [**k-u** **kaen-en** n-ira?]  
NM what NOM-NM- eat-UV GEN.3SG  
 ‘What did he eat?’ (lit. ‘It is a what the x that was eaten by him?’)
- b. U turun [**k-u** **kaen-en** n-ira].  
NM glutinous.rice NOM-NM eat-UV ~~NOM.3SG~~  
 ‘It’s glutinous rice (that) he ate.’ (lit. ‘It’s *turun* that was eaten by him.’)
- c. **Kaen-en**=tu n-ira k-iyā turun.  
eat-UV=PFT GEN-3SG NOM-ANAPH glutinous.rice  
 ‘He’s eaten that glutinous rice.’
- d. \*U maan k-u **mi-kaen** cira?  
NM what NM-NOM AV-eat NOM 3SG  
 Intended: ‘What did he eat?’

In (54b,c), the undergoer role of *u maan* ‘what?’ and *u mi-takaw-ay* ‘thief’ is indexed by the locative voice *libabuy-an* in the RC. Compare with the declarative sentence (54a).

- (54) a. Libabuy-**an** n-*iya* wacu **k-u** mi-takaw-ay.  
 bark-LOC.V GEN-ANAPH dog NOM-NM AV-steal-NMZ  
 ‘The dog is barking at a thief.’ (lit. ‘A thief is barked at by the dog.’)
- b. [U maan] k-u libabuy-**an** n-*iya* wacu?  
 NM what NOM-NM bark-LOC.V GEN-ANAPH dog  
 ‘What is this dog barking at?’ (lit. ‘It’s a what that’s being barked at by the dog?’)
- c. [U mi-takaw-ay] k-u libabuy-**an** n-*iya* wacu.  
 NM AV-steal-NMZ NOM-NM bark-LOC.V GEN-ANAPH dog  
 ‘It’s a thief that the dog is barking at.’

In (55a,b), the instrument role of *u maan* ‘what?’ and *u qepah* ‘wine’ is indexed by *sa-pa-kaen*.<sup>22</sup> Compare with the declarative sentence in (55c). And compare also with an equative, topic-comment construction (*u* NP, *u* NP, see 4.1) such as (55d), in which the whole *u* marked substantival clause is the topic, optionally marked by *i(ri)*, while *u qepah* is the comment/predicate.

- (55) a. U maan [k-u sa-pa-kaen=isu t-u cabay?].  
 NM what NOM-NM INST.V-CAUS-eat=GEN.2SG OBL-NM friend  
 ‘What did you give (your) friend to drink? (lit. ‘It is a what the x used to give to drink by you to the friend?’)
- b. U qepah [k-u sa-pa-kaen numaku<sup>23</sup> t-u cabay].  
 NM wine NOM-NM INST.V-CAUS-eat GEN.1SG OBL-NM friend  
 ‘It’s wine that I gave to drink to my friend.’
- c. Sa-pa-kaen=aku t-u cabay=aku k-u qepah.<sup>24</sup>  
 INST.V-CAUS-eat=GEN.1SG OBL-NM friend=GEN.1SG NOM-NM wine  
 ‘I gave my friend wine to drink.’
- d. [U sa-pa-kaen=aku t-u cabay i]/ u qepah.  
 NM INST.V-CAUS-eat=GEN.1SG OBL-NM friend TOP NM wine  
 ‘As for the drink I gave to (my) friend, it is water.’

**6.3.2 Cleft narrow foci: core arguments.** Cleft core NP pivots trigger a nominative gerund RC ([*k-u* RC]) that contains the presupposition. Their semantic role is indexed by verbal voice in the RC. The case of cleft nonpivot, peripheral arguments, and cleft adjuncts is analyzed in 6.4.2, 6.5, and 6.6.

**6.3.2.1 Cleft informative foci in Actor role with *cima* ‘who?’ and *maan* ‘what?’.** *Cima* ‘who?’ is a case-marked interrogative pronoun, with no verbal derivation, while *maan* ‘what?’ is either a pronoun (*u maan* ‘a what?’), or a derived verb (see 6.2). Their focus domains are thus different: *maan* allows argument, predicate, and sentence foci, while *cima* is restricted to narrow argument focus.

22. *Kaen* [kaʔən] ‘eat’ includes ingesting liquids other than water (for example, alcohol, soup).

23. The clitic =*aku* or the full form *numaku* occur in distinct speech levels; full forms occur in careful speech.

24. The reverse order is rejected: \**Sa-pa-kaen=aku k-u qepah tu cabay=aku*.

**Narrow focus with *cima* in Actor role.** *Cima* ‘who?’ derives from Proto-Austronesian \**si-ima*<sup>25</sup> ‘who?’ (Blust 2006); the form *cima* incorporates the *i* article for personal nouns, accounting for its incompatibility with *u* (\**u cima*), unlike *u maan* ‘a what?’. *Cima* ‘who?’ is the nominative form, *cima-an* ‘to whom?’ is oblique, *n-ima* is genitive and marks either possession ‘whose?’ or the agent of a transitive construction.

*Cima* must be in sentence-initial position, as in (56b), the *in situ* position (the position of the variable, \**mi-senat cima*?) was rejected by speakers. In (56b), *cima* is predicative, marked for past tense, and its actor role is indexed on the gerund verb *mi-senat-ay*<sup>26</sup> in the nominative RC. Compare with the UV construction in (56a). The answer (56c) shows a stressed and lengthened cleft free pronoun *aku* [aku:] marked for perfect (by =*tu*). An accidental, self-inflicted wound requires a different verb, *ma-tabal*.

- (56) a. Ma-senat k-u taruduq n-u tireng.  
 UV-cut NOM-NM finger GEN-NM self/body  
 ‘My own finger was cut.’
- b. Na [cima] [k-u mi-senat-ay t-u taruduq=isu?]  
 PST who NOM-NM AV-cut-NMZ OBL-NM finger=GEN 2SG  
 < focus > <————— presupposition —————> (IS level)  
 < predicate >————— gerund RC argument —————> (syntactic level)  
 ‘Who cut your finger?’
- c. [Aku:=tu] [k-u ma-tabal-ay t-u taruduq n-u tireng.]  
 FR 1SG=PFT NOM-NM NAV-wound-NMZ OBL-NM finger GEN-NM self/body  
 ‘I wounded my own finger.’ (lit. ‘It was me the wounder of my finger.’)

**Narrow focus with *u maan* in Actor role.** In (57a), the actor role of the cleft WH *u maan* is indexed on the gerund verb *k<em>an-ay*, which has an oblique argument *tu tatakulaq*. Compare with the declarative sentence (57b).

- (57) a. Na u maan k-u k<em>an-ay t-u tatakulaq?  
 PST NM what NOM-NM <NAV>eat-NMZ OBL-NM frog  
 ‘What is it that ate the frog?’
- b. K<em>an t-u tatakulaq k-u quner.  
 <NAV>eat OBL-NM frog NOM-NM snake  
 ‘The snake ate the frog.’

Motion verbs are NAV *ma*-verbs like *ma-rebahui* (without the \**mi*-form) (58a). Both clefts, *u maan* and *u paypay*, in (58b,c) are nonverbal predicates, with nonfinite nominative RCs.<sup>27</sup> Their semantic role as entities in motion<sup>28</sup> is indexed on the gerund verb *ma-rebahuy-ay*.

- (58) a. Ma-rebahui k-u ayam i kudukudul.  
 NAV-fly NOM-NM bird LOC RED.above  
 ‘A bird’s flying in the sky.’

25. Proto-Austronesian \**si* is the nominative case for singular personal nouns and \**ima* meant ‘who?’.

26. Compare *ma-tabal* ‘cut o.s.’ accidentally, with *mi-senat* ‘cut’ (intentionally).

27. Wu (2006) analyses these constructions as headless relative clauses.

28. Motion verbs are NAV *ma*-verbs and behave as verbs expressing a type of bodily function.

- b. **U maan k-iya ma-rebahuy-ay i kudukudul, u ayam haw?**  
 NM what NOM-ANAPH NAV-fly-NMZ LOC RED.above NM bird QM  
 ‘What’s flying in the sky, is it a bird?’ (lit. ‘It’s a what that “flier” ...’)
- c. **U paypay k-iya ma-rebahuy-ay.**  
 NM kite NOM-ANAPH NAV-fly-NMZ  
 ‘It’s a kite that’s flying.’

**6.3.2.2 Narrow contrastive foci: cleft Actors.** Contrastive cleft foci with *haw~saw* and informative cleft foci (answering WH-questions) behave similarly. The cleft form is predicative, possibly marked for tense-aspect (59a), and can be negated (59b).

- (59) a. **Inacila [na isu haw] k-u tayni-ay?**<sup>29</sup>  
 yesterday PST FR.2SG QM NOM-NM come-NMZ  
 < focus/predicate > < clausal argument >  
 ‘Was it you who came yesterday?’ (lit. ‘Was it you the arriver ...?’)
- b. **Caay ka aku, [u saba=aku] k-u tayni-ay.**  
 NEG NFIN FR.1SG NM young.sibling=GEN.1SG NOM-NM come-NMZ  
 ‘No it was not me, it was *my younger brother* who came.’

Compare the declarative sentence (60a) with the cleft in (60b), where the cleft noun *u matuasay* ‘parents’ is marked for past tense, and its Actor role is indexed by AV *mi-* on the gerund verb *mi-diput-ay* in the RC.

- (60) a. **Na mi-diput k-u ma-tuas-ay itakuwan(an).**  
 PST AV-protect NOM-NM NAV-old-NMZ OBL.1SG(OBL)  
 ‘(My) parents brought me up.’
- b. **Na u ma-tuas-ay k-u mi-diput-ay itakuwan(an).**  
 PST NM NAV-old-NMZ NOM-NM AV-protect-NMZ OBL.1SG(OBL)  
 ‘It was older people who brought me up.’  
 (urip nu Balah.007/Balah’s life)

**6.3.2.3 Narrow focus: cleft Undergoer.** Moving on to cleft undergoers, the Undergoer role is indexed on the UV *-en* verb form in the RC, as in (61).

- (61) **Anini a uraan k-u kungku-en=aku t-u seikacu n-u Pangcah.**  
 now FUT that NOM-NM narrate-UV=GEN.1SG OBL-NM customs GEN-NM Amis  
 ‘Now this is what I’m going to tell about the Amis’s customs.’ (lit. ‘It is that what is going to be told by me ...’) (Arikakai.015)

In (62), the Undergoer role of the cleft *u lusay* is indexed by the locative *-an* on the perfect nominalized verb form in the RC.

- (62) **U lusay [k-u ni-peruq-an n-i ama].**  
 NM fruit NOM-NM PFT NMZ-pick-LOC GEN-PM father  
 ‘It’s fruit that was picked by father.’

Compare with the relative clause in (63): the head *lusay* is modified by the prenominal RC marked at its right edge by the linker *a*. In the relative clause (63), as in the cleft construction (62), the undergoer role of *lusay* is indexed by the locative *-an*.

29. *Tayni* is a deictic Ø-motion verb.

- (63) Caay ka-alsuq [k-u [ni-peruq-an n-i ama a] lusay].  
 NEG NFIN-tasty NOM-NM PFT NMZ-pick-LOC GEN-PM father LNK fruit  
 ‘The fruit that was picked by father is bad.’

**6.3.2.4 Narrow focus: cleft Instrument.** In contrastive or informative cleft constructions alike, cleft core arguments with instrument role are indexed by *sa-* prefixed to nonfinite verb forms as *sa-pi-* or *sa-ka-*.<sup>30</sup>

**Contrastive clefts.** In (64a,b), the cleft instrument (questioned by *haw~saw* in [64a]) is indexed on the gerund form *sa-pi-*, which retains an oblique argument.

- (64) a. [Na unian haw a hawan] k-u sa-pi-cikcik  
 PST DX1 QM LNK machete NOM-NM INST.V-NFIN.AV-cut  
 numisu t-u titi?  
 GEN.2SG OBL-NM meat  
 ‘Was it with this machete that you cut the meat?’ (lit. ‘Was it this machete the (thing) you used to cut the meat with?’)
- b. Caay! [uraan a hawan] k-u sa-pi-cikcik t-ina titi.  
 no DX3 LNK machete NOM-NM INST.V-NFIN AV-cut OBL-DX2 meat  
 ‘No, it’s that machete that I used to cut the meat with.’ (lit. ‘It is that machete the (thing) I used to cut the meat.’)

In (65), the cleft instrument *u sakubad nu ayam* ‘bird’s wings’ is indexed on the nonfinite verb form *sa-ka-rebahui*.

- (65) U sakubad n-u ayam [k-u sa-ka-rebahui] a tayra  
 NM wing GEN-NM bird NOM-NM INST.V-NFIN-fly CMP go.there  
 < — predicative focus — > < argument/presupposition >  
 i ?a-?adup-an.  
 LOC CA RED-hunt-LOC  
 ‘It’s bird’s wings that were used to fly to arrive to the hunting place.’  
 (Lalangawan.020)

**Informative cleft instrument.** The instrument role of the clefts *u maan* and *u puut* are similarly indexed by *sa-* (66a,b). *U maan* must be cleft to sentence-initial position as the syntactic pivot, the *in situ* position, is rejected.

- (66) a. U maan k-u sa-pi-cikcik=i<sup>31</sup> t-u dateng?  
 NM what NOM-NM INST.V-NFIN AV-cut=GEN 2SG OBL-NM vegetables  
 ‘What did you use to cut the vegetables/what did you cut the vegetables with?’
- b. U puut k-u sa-pi-cikcik=aku t-u dateng.  
 NM knife NOM-NM INST.V-NFIN.AV-cut=GEN 1SG OBL-NM vegetables  
 ‘It’s a knife that I used to cut the vegetables/I used a knife to ...’

**6.4 CLEFT OBLIQUE ARGUMENTS.** As shown in 6.3, cleft core arguments of AV *mi-*verbs and of extended intransitive NAV *ma-*verbs, as well as cleft undergoers of transitive UV *ma-*verbs, must be the syntactic pivot, with their semantic role indexed by the corresponding voice of the verb in the RC. The analysis now centers on the *t-u-*

30. In derived forms, *mi-*verbs occur as nonfinite *pi-V* forms, *ma-*verbs occur as nonfinite *ka-V* forms.

31. Compare with the declarative sentence: *Mi-cikcik k-aku t-u dateng* ‘I’m chopping the vegetables’.



marked themes of extended intransitive *mi-* and *ma-*verbs and shows some differential treatment as pivots (6.4.1) or as nonpivot/peripheral arguments (6.4.2) varying with subtypes of verbs (psych and discourse) and with the degree of affectedness of the theme.

**6.4.1 Cleft themes of extended intransitive *mi-* and *ma-*verbs.** Extended intransitive AV *mi-*verbs and NAV *ma-*verbs<sup>32</sup> have a nominative pivot and a *t-(u)*-marked theme. These theme arguments are generally cleft as syntactic pivots, and are indexed by a UV *-en* or a locative *-an* verb form in the *k-u* RC.

**Cleft themes of extended intransitive AV *mi-*verbs.** Compare the declarative sentence in (67a) with the constructions in (67b,c), where *u maan* (67b) and *u tatakulaq* (67c) are cleft with their semantic role indexed by the UV *-en* form in the RC (Ca-reduplication marks atelic aspect).

- (67) a. Mi-kilim cira t-iyā tatakulaq-an n-ira.  
 AV-look.for NOM.3SG OBL-ANAPH frog-OBL GEN-3SG  
 'He's looking for his frog.' (Frog story.078)
- b. U maan k-u ka-kilim-en=isu?  
 NM what NOM-NM CA.RED-look.for-UV=GEN.2SG  
 'What are you looking for?' (lit. 'Who is the one looked for by you?')
- c. U tatakulaq k-u ka-kilim-en=aku.  
 NM frog NOM-NM CA RED-look.for-UV=GEN.1SG  
 'It's a frog that I'm looking for.'

**Cleft oblique arguments of extended intransitive *ma-* verbs.** Cleft themes of extended intransitive NAV *ma-*verbs are syntactic pivots, whose role is indexed by the locative *-an* verb form in the RC clause. The syntactic behavior is similar in contrastive foci with the question marker *haw~saw* (68b,c) or in informative WH-foci (69b). Compare the declarative clauses in (68a, 69a) with the cleft theme arguments in (68b,c) and (69b).

- (68) a. Ma-ngalay k-aku t-ina kabang-an.  
 NAV-want NOM-1SG OBL-DX2 bag-OBL  
 'I'd like this bag.'
- b. Uraan haw k-u ka-ngalay-an=isu?<sup>33</sup>  
 that QM NOM-NM NFIN-want-LOC=GEN.2SG  
 'Is it that one that you want?'
- c. Uraan k-u ka-ngalay-an=aku.  
 that NOM-NM NFIN-want-LOC=GEN.1SG  
 'It's *that one* that I want.'

**Differential marking of themes: cleft vs. *in situ* WH-constructions.** Cleft constructions with the WH-word standing as the syntactic pivot are obligatory if new information is requested as in (69b) and (70b). But if the question is confirmative or requests repetition of old information, and with verbs whose themes are superficially affected, like *mi-cihi* 'scold', or *ma-*experiencer or psych verbs, the oblique WH-form may remain *in situ* and is then treated as a nonpivot, oblique argument, as in (69c) where *t-u maan*, which indexes the source of the feeling, may remain *in situ*.

32. While transitive UV *ma-*verbs have a nominative Undergoer and a genitive agent.

33. *Pi-...-an* and *ka-...-an* are nonfinite locative verb forms.

- (69) a. Ma-talaw k-uhni t-u kawas.  
 NAV-be.scared NOM.3PL OBL-NM spirit  
 ‘They’re scared of spirits.’
- b. U maan k-u ka-talaw-an n-ira?  
 NM what NOM-NM NFIN-fear-LOC GEN-3SG  
 ‘What is he scared of?’ (lit. ‘About what is his fright?’)
- c. Ma-talaw k-uhni t-u maan?  
 NAV-be.scared NOM.3PL OBL-NM what  
 ‘They’re scared of what?’

Similarly, while *cima* is the syntactic pivot of a cleft construction in (70b), the oblique WH-form *cimaan* may remain *in situ* in (70c) if the question is confirmative.

- (70) a. Mi-cihi t-u wawa k-u ma-tuas-ay.  
 AV-scold OBL-NM child NOM.NM NAV-old-NMZ  
 ‘The old person is/was yelling/yelled at the child.’
- b. Cima k-u cihi-en n-i Balah?  
 who NOM-NM scold-UV GEN-PM Balah  
 ‘Who did Balah scold?’ (lit. ‘Who was scolded by Balah?’)
- c. Mi-cihi ci Balah cima-an?  
 AV-scold NOM PM Balah who-LOC  
 ‘Balah yelled at whom?’

**6.4.2 Cleft peripheral oblique arguments: *a* CMP clauses.** The more peripheral oblique arguments now discussed are (i) the superficially affected oblique themes of a subtype of AV *mi*-verbs such as ‘pay a visit’ or ‘yell at’, and (ii) prepositional arguments, including those of three-place *mi*-verbs expressing the source. When cleft, these trigger *a*-marked complement clauses.

**6.4.2.1 Differential marking of superficially affected themes of AV *mi*-verbs.**

Depending on aspect, some AV *mi*-verbs such as *mi-lisuiq* ‘pay a visit’ and *mi-cihi* ‘yell at’ show differential treatment of their theme argument, with consequences on alignment when the theme is cleft. Consider (71a), repeated from (70b): in completed aspect (71a), the theme is treated as an affected undergoer and, when cleft, is the syntactic pivot of a UV *-en* verb form in the RC. In ongoing aspect (71b), the superficially affected theme is treated as a peripheral argument, it is cleft with its oblique case-marking, and it triggers a finite *a*-marked complement clauses.<sup>34</sup>

- (71) a. Cima k-u cihi-en n-i Balah?  
 who NOM-NM scold-UV GEN-PM Balah  
 ‘Who did Balah scold?’ (lit. ‘Who was scolded by Balah?’)
- b. Cima-an a mi-cihi ci Balah?  
 who-LOC CMP AV-scold NOM.PM Balah  
 ‘At whom is Balah yelling?’

Cleft oblique prepositional arguments of verbs like *mi-lisuiq* also trigger *a* CMP clauses.

34. But \**Cima k-u mi-cihi ci Balah* with the intended meaning ‘Who did Balah scold?’ is ungrammatical, as the role of *cima* would stand in conflict with AV voice.

- (72) a. Mi-lisuq ci Dihang (i) cira-an.  
 AV-visit NOM-PM Dihang LOC 3SG-OBL  
 ‘Dihang is paying a visit to him/her.’
- b. (I) cima-an a mi-lisuq ci Dihang?  
 LOC who-OBL CMP AV-visit NOM PM Dihang  
 ‘Who is Dihang paying a visit to?’

#### 6.4.2.2 Differential marking of locative arguments of three-place AV *mi*-verbs.

Similarly, the cleft prepositional arguments of three-place AV *mi*-verbs expressing the source (marked by locative *i*) display differential treatment as pivots (73b) or as peripheral arguments (73c). Compare the declarative clause (73a) with (73b), where *cima* is the pivot and its source role is indexed by locative *-an* in the gerund RC. On the other hand, in (73c,d), the cleft peripheral prepositional arguments *i cima-an* and *i ci Dungi-an* trigger a finite *a* CMP clause with a *mi*-verb. In (73c), the Actor pivot *k-isu* is raised from the CMP clause as the argument of the cleft predicate. Raising is optional, as shown by the position of *k-aku* in (73d). The *in situ* position of the prepositional WH-word is again restricted to confirmative questions (73e) and acceptability judgments are not unanimous among speakers.

- (73) a. Mi-caliw k-aku t-u pida i ci Dungi-an.  
 AV-borrow NOM-1SG OBL-NM money LOC PM Dungi-OBL  
 ‘I borrowed money from Dungi.’
- b. Na cima k-u pi-caliw-an=isu t-u pida?  
 PST who NOM-NM NFIN.AV-borrow-LOC=GEN.2SG OBL-NM money  
 ‘Who did you borrow the money from?’ (lit. ‘Who is the one who was borrowed money by you?’)
- c. Na i cima-an k-isu a mi-caliw t-u pida?  
 PST LOC who-OBL NOM-2SG CMP AV-borrow OBL-NM money  
 ‘To whom did you borrow the money?’
- d. Na i ci Dungi-an a mi-caliw k-aku (t-u pida).  
 PST LOC PM Dungi-OBL CMP AV-borrow NOM-1SG OBL-NM money  
 ‘It’s from Dungi that I borrowed (the money).’
- e. Na mi-caliw k-isu t-u baru i cima-an?  
 PST AV-borrow NOM-2SG OBL-NM flower LOC who-OBL  
 ‘You bought the flowers from whom?’

**6.5 CLEFT AGENTS.** Agents are not indexed by the verbal voice system; they are nonpivot arguments occurring in transitive UV *ma*-constructions, in passive UV *-en* constructions (74a), and in all other voices (instrumental, locative). As expected with cleft nonpivot arguments, agent extraction triggers an *a* CMP clause as in (74b)<sup>35</sup> with a nonfinite verb, marked by the perfective nominalizing prefix *ni-*, while the nominative undergoer pivot *kira wawa* is indexed by the locative form *-an*. A finite UV verb form *cih(e)n* is disallowed in a cleft like (74b) (\**Na u n-ima a cih(e)n k-iy a wawa?*). Being nonpivots, the genitive agent WH- forms *nu maan* (non-humans) and *nima* (humans)

35. Compare with possessive determination: *u n-ima a cudad?* (NM GEN-who LNK book) ‘whose book is it?’.

may remain *in situ* but only if confirmative (74c,d). They are otherwise cleft to predicate initial position (74b).

- (74) a. **Cihi-(e)n** n-u wina k-iyá wawa.  
 scold-UV GEN-NM mother NOM-ANAPH child  
 ‘The child was scolded by mother.’
- b. Na u **n-ima** a **ni-cihi-an** k-ira wawa?  
 PST NM GEN-who CMP PFT NMZ-scold-LOC NM-DX3 child  
 ‘Who was it that the child was scolded by?’
- c. **Cihi-(e)n** **n-ima** k-iyá wawa?  
 scold-UV GEN-who NM-ANAPH child  
 ‘The child was scolded by whom?’
- d. Ma-kaen **n-u** **maan** k-u tatakulaq?  
 UV-eat GEN-NM what NOM-NM frog  
 ‘The frog was eaten by what?’

**6.6 CLEFT PLACE AND TIME ADJUNCTS.** Cleft place and time adjuncts expectedly trigger a CMP clauses without voice indexing. In (75a), the cleft prepositional adjunct *i Tayliku* is predicative and marked for past tense. The raising of the Actor pivot (*k-isu*) of the AV *mi*-verb from the complement clause contributes to the clausal integration of a cleft adjunct heading a complement clause, barring its interpretation as a merely fronted adjunct. It also corresponds to an optional tendency for nominative pivot pronouns to raise as arguments of the cleft predicate: see (73c), (75a). Compare with the declarative sentence (75b).

- (75) a. **Na i** Tayliku **haw** k-isu a mi-nanam t-u suwal  
 PST LOC China QM NOM-2SG CMP AV-know OBL-NM language  
 n-u Hulam?  
 GEN-NM Chinese  
 ‘Was it in mainland China that you learned the Chinese language?’
- b. Mi-nanam **k-aku** t-u suwal n-u Hulam **i** Tayliku.  
 AV-know NOM-1SG OBL-NM language GEN-NM Chinese LOC China  
 ‘I learned Chinese in mainland China.’

Locative *icuwa* ‘where?’ behaves similarly. As a cleft WH-word in predicate function, it heads a complement clause with a raised pronoun (76).

- (76) Icuwa k-isu a m-aruaq?  
 where NOM-2SG CMP NAV-stay  
 ‘Where do you live?’

Like *maan* ‘what?’, *icuwa* ‘where?’ can be derived as a voice-affixed verb, restricted to the passive UV *-en* form<sup>36</sup> with an undergoer pivot (77); in that case, no cleft is involved: it is a one-clause interrogative construction vs. the two-clause cleft construction in (76).

- (77) A icuwa-**en**=aku k-ina cudad?  
 FUT where-UV=GEN.2SG NOM-DX2 book  
 ‘Where shall I (put) this book?’

36. *Icuwa* ‘where?’ does not allow *mi*- or *ma*, nor does *hacuwa* ‘when?’.

Cleft time adjuncts, WH-form *hacuwa* ‘when?’, or lexical adjuncts, must also be cleft, triggering a CMP clauses (78). *Hacuwa* is a predicate marked for future; it must be cleft—\**Tayni kisu hacuwa?* is ungrammatical due to the verb-like properties of *hacuwa*.

- (78) A *hacuwa* k-isu a *tayni?*  
 FUT when NOM-2SG CMP arrive  
 ‘When will you arrive?’

On the other hand, when *hacuwa* is a unit of count meaning ‘how much?’, the gerund [*k-u* RC] construction is available. *Hacuwa* is predicative and must be in sentence-initial position.

- (79) *Hacuwa*=tu [**k-u ni-but-ing-an** numisu] *tangasa anini?*  
 how.much=PFT NOM-NM PFT.NMZ-fish-LOC GEN.2SG reach now  
 ‘How much fish have you caught until now?’ (lit. ‘It’s how much your catch of fish?’)

**6.7 SUMMARY.** To sum up, both core and peripheral arguments may be cleft, but with radically different constructions. Core arguments are cleft as syntactic pivots and their semantic role is indexed by voice-marked verbs (as Actor, Undergoer, Instrument, Locus) in the gerund [*k-u* RC] construction containing the presupposition. Cleft noncore arguments, and adjuncts, trigger a CMP clauses. Noncore arguments include (i) agents of transitive UV *ma*-verbs (6.5), (ii) superficially affected oblique themes of some AV *mi*-verbs, (iii) prepositional arguments of some *mi*-verbs and of three-place AV *mi*-verbs, expressing destination or source (6.4.2). Still, in the last two cases, oblique themes may be treated as the pivots of UV *-en* or LOC *-an* verb forms.

Another central feature is that WH-forms requesting new information must be cleft; few WH-forms may remain *in situ*. When they do, they only request confirmation of old information; the *in situ* position of WH-forms is thus pragmatically driven. It is also restricted to the same nonpivot arguments mentioned above and to oblique arguments of extended intransitive NAV *ma*-psych verbs (6.4.1).

**6.8 THE SYNTAX OF CLEFT FOCI WITHIN SUBORDINATE CLAUSES.** I now turn to cleft constructions within complement clauses, such as (80a). Pivot arguments (such as *Balah*) can be extracted and cleft to the initial position of an asyndetic complement clause, but they do not move further left beyond this position; they remain in the position of the antecedent of the following gerund RC. In (80a), the Actor role of *ci Balah* is indexed on the gerund AV verb form *mi-...-ay*. *Ci Balah* is predicative, marked for past tense; the gerund *k-u* RC is the nominative argument. Compare this with the noncleft construction in (80b).

- (80) a. Ma-banaq k-aku [na ci Balah [**k-u mi-araw-ay** t-iyā  
 NAV-know NOM-1SG PST PM Balah NOM-NM AV-see-NMZ OBL-ANAPH  
 < focus/predicate > — gerund RC, presupposition —  
 tamdaw-an]].  
 person-OBL  
 —————>  
 ‘I know it was Balah who met that person.’

- b. Ma-banaq k-aku [mi-araw ci Balah t-iyā tamdaw-an].  
 NAV-know NOM-1SG AV-see PM Balah OBL-ANAPH person-OBL  
 ‘I know that Balah met that person.’

In (81), the source role of cleft *cima*, that is, the seller, is indexed by the gerund locative form *ca-cakay-an* in the RC.

- (81) Awaay k-u ma-banaq-ay [[anu **cima**] [k-u ca-cakay-an  
 NEG.EXS NOM-NM NAV-know-NMZ perhaps who NOM-NM CA.RED-trade-LOC  
 n-ira t-u karireng]].  
 GEN-3SG OBL-NM car  
 ‘No one knows (i.e., there’s no knower) who he may have bought the car from.’

In (82), the cleft undergoer *u maan* is indexed by the locative form on the perfective nominalized verb form *ni-ka-melaw-an* in the RC. *U maan* may not occur *in situ* (no confirmative reading being available), it must be cleft.

- (82) Ma-ngalay k-aku a ma-banaq [[**u maan**]  
 NAV-want NOM-1SG CMP NAV-know NM what  
 [k-u ni-ka-melaw-an n-i Dungi]].  
 NOM-NM PFT.NMZ-NFIN-see-LOC GEN-PM Dungi  
 ‘I want to know what was seen by Dungi.’

Compare with a complement clause in which the interrogative verb *mi-maan* ‘do what?’ operates sentence/event focus (83). Also note the optional raising of *k-aku* after the negation, rather than after the negated verb *caay ka-banaq kaku a mi-maan kita...*

- (83) Caay k-aku ka-banaq [a **mi-maan** k-ita].  
 NEG NOM-1SG NFIN-know FUT AV-do.what NOM-1PL.INCL  
 ‘I do not know what we are going to do.’ (Kasubucan.019)

**7. MAIN RESULTS AND CONCLUSION.** This discussion of information structure and of the basic voice system of Northern Amis has shown that they belong to distinct but interacting levels. All arguments and adjuncts can be left-dislocated as topics, independently from any syntactic pivot constraint in the voice system; they are bounded by a pause, and optionally by *i(ri)*. Topics must be minimally specific, while brand new entities and indefinite, nonspecific nouns tend to be introduced by existential constructions. Yet, in contrast with Tagalog, indefinite nouns may occur as the nominative pivots of finite verbs. Thus, definiteness has a limited impact on the choice of AV *mi-* and *UV ma-* forms. Shuanfan Huang reached a similar conclusion about Seediq (Formosan): “although some AV clauses take non-referential patients, lexical patient pivots in NAV are not significantly more referential and/or definite than those in AV clauses” (2002:675).

Evidential foci do not trigger any syntactic restructuring: they are *in situ* foci under the scope of *han(=tu)* and *sa(=tu)* and signal “alleged.” second-hand information or personal judgment. Just like left-dislocated topics, evidential foci do not fall under illocutionary scope and are not questioned, nor negated: they only occur in declarative clauses with epistemic stance. Cleft foci, on the other hand, are questioned and negated, like all other predicates. Besides, cleft narrow foci (informative or restrictive) are quite constrained: a

cleft core argument must be the syntactic pivot of the verb and its semantic role must be indexed by verbal voice in the [*k-u* gerund RC] containing the presupposition. No such restriction applies to agents (which are nonpivots), nor to cleft adjuncts: those trigger *a* CMP clauses. When cleft, the prepositional arguments of three-place AV *mi*-verbs and the superficially affected themes of some *mi*-verbs display differential marking depending on their treatment as core or as peripheral arguments, respectively triggering [*k-u* gerund RC] or *a* CMP clause.

## REFERENCES

- Aikhenvald, Alexandra Y. 2004. *Evidentiality*. Oxford: Oxford University Press.
- Aikhenvald, Alexandra Y., and R. M. W. Dixon. 1998. Dependencies between grammatical systems. *Language* 74:56–80.
- Blust, Robert. 1999. Subgrouping, circularity and extinction: Some issues in Austronesian comparative linguistics. In *Selected papers from the Eighth International Conference on Austronesian Linguistics*, ed. by Elizabeth Zeitoun and Paul Jen-kuei Li, 31–94. Taipei: Institute of Linguistics, Academia Sinica.
- Bril, Isabelle. 2007 [2011]. Coordination, information hierarchy and subordination in some Austronesian languages. In *Converbs, medial verbs, clause chaining and related issues*, ed. by Azeb Amha, Christian J. Rapold, Sascha Völlmin, and Silvia Zaugg-Coretti. *Frankfurter Afrikanistische Blätter* 19:213–43. Rüdiger Köppe Verlag.
- . 2010. Informational and referential hierarchy: clause-linking strategies in Austronesian-Oceanic languages. In *Clause-linking and clause hierarchy: Syntax and pragmatics*, ed. by Isabelle Bril, 269–311. (Studies in Language Companion Series 121). Amsterdam: Benjamins.
- Chen, Teresa. 1987. *Verbal constructions and verbal classifications in Nataoran-Amis*. Canberra: Pacific Linguistics.
- Edmonson, Jerold, John Esling, Jimmy Harris, and Tung-chiou Huang. 2005. A laryngoscopic study of glottal and epiglottal/pharyngeal stop and continuant articulations in Amis, an Austronesian language of Taiwan. *Language and Linguistics* 6(3):381–96.
- Erteschik-Shir, Nomi. 2014. Information structure and (in)definiteness. In *Cross-linguistic studies on noun phrase structure and reference*, ed. by Patricia Cabredo Hofherr and Anne Zribi-Hertz, 23–51. *Syntax and Semantics* 39. Leiden: Brill.
- Foley, William A., and Robert D. van Valin. 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.
- Gundel, Jeanette K., and Nancy Hedberg. 2015. Reference and cognitive status: Scalar inference and typology. In *Information structuring of spoken language from a cross-linguistic perspective*, ed. by M. M. Jocelyne Fernandez-Vest and Robert D. van Valin, 33–54. Berlin/Boston: Mouton de Gruyter.
- Guentchéva, Zlatka (ed.). 1996. *L'énonciation médiatisée*. Louvain/Paris: Peeters.
- Himmelman, Nikolaus P. 2005. The Austronesian languages of Asia and Madagascar: Typological characteristics. In *The Austronesian languages of Asia and Madagascar*, ed. by Alexander Adelaar and Nikolaus P. Himmelman, 110–81. London: Routledge.
- Huang, Lillian M. 1995. The case markers and pronominal system in Amis. *The Journal of National Chengchi University* 70:217–58.
- Huang, Shuanfan. 2002. The pragmatics of focus in Tsou and Seediq. *Language and Linguistics* 3(4):665–94. Taipei: Academia Sinica.

- Kaufman, Daniel. 2009. Austronesian typology and the nominalist hypothesis. In *Austronesian historical linguistics and culture history: A festschrift for Robert Blust*, ed. by Alexander Adelaar and Andrew Pawley, 197–226. Canberra: Pacific Linguistics.
- Keenan, Edward, and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8:63–99.
- . 1979. Data on the noun phrase accessibility hierarchy. *Language* 55:333–51.
- Krifka, Manfred. 2007. Basic notions of information structure. In *The notions of information structure*, ed. by Caroline Féry, Gisberg Fanselow, and Manfred Krifka, 13–55. Potsdam: Universitätsverlag.
- Krifka, Manfred, and Renate Musan (ed.). 2012. *The expression of information structure*. Berlin/Boston: Mouton De Gruyter.
- Kroeger, Paul. 1993. *Phrase structure and grammatical relations in Tagalog*. Stanford: CSLI Publications.
- Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, Focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.
- . 2000. When subjects behave like objects: a markedness analysis of sentence focus constructions across languages. *Studies in Language* 24:611–82.
- Lazard, Gilbert. 2001. On the grammaticalization of evidentiality. *Journal of Pragmatics* 33:358–68.
- Liu, Dorinda Tsai-hsiu. 1999. Cleft constructions in Amis. MA thesis, National Taiwan University.
- Ross, Malcolm. 2009. Proto Austronesian verbal morphology: A reappraisal. In *Austronesian historical linguistics and culture history. A festschrift for Robert Blust*, ed. by Alexander Adelaar and Andrew Pawley, 285–316. Canberra: Pacific Linguistics.
- Sagart, Laurent. 2004. The higher phylogeny of Austronesian and the position of Tai-Kadai. *Oceanic Linguistics* 43:411–44.
- Shibatani, Masayoshi. 2009. Elements of complex structures, where recursion isn't: The case of relativization. In *Syntactic complexity. Diachrony, acquisition, neuro-cognition, evolution*, ed. by T. Givón and Masayoshi Shibatani, 163–98. Typological Studies in Language 85. Amsterdam: Benjamins.
- Starosta, Stanley. 1995. A grammatical subgrouping of Formosan languages. In *Austronesian studies relating to Taiwan*, ed. by Paul Jen-kuei Li, Cheng-hwa Tsang, Yingkuei Huang, Dah-an Ho, and Chiu-yu Tseng, 683–726. Taipei: Institute of History and Philology, Academia Sinica.
- Tsuchida, Shigeru. 1988. Amis. *The Sanseido encyclopedia of linguistics*, vol. 1: *Languages of the world, Part one*, ed. by Takashi Kamei, Rokuro Kono, and Eiichi Chino, 447–49. Tokyo: Sanseido.
- Tsukida, Naomi. 2008. Verb classification in Amis. In *The typology of semantic alignment*, ed. by Mark Donohue and Søren Wichmann, 277–93. Oxford: Oxford University Press.
- Van Valin, Robert D. 2005. *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press.
- Wu, Joy. 2006. Verb classification, case marking, and grammatical relations in Amis. PhD diss., State University of New York at Buffalo.
- Zeitoun, Elizabeth, Lillian M. Huang, Marie M. Yeh, and Anna H. Chang. 1999. Existential, possessive, and locative constructions in Formosan languages. *Oceanic Linguistics* 38:1–42.