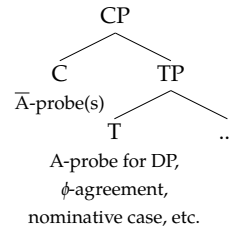


Extraction and licensing in Toba Batak

Michael Yoshitaka ERLEWINE, National University of Singapore, mitcho@nus.edu.sg
GLOW 40, March 2017

The plot

- (1) **Traditional division of labor:** (Chomsky, 1986, a.o.)
 a. C: \bar{A} -probe(s) for certain XP(s) (focus, *wh*, ...)
 In some languages, Spec,CP must be filled (V2)
 b. T: A-probe for DP
 ϕ -agreement, case assignment (nominative)
 in some languages, Spec,TP must be filled (EPP)



- (2) **Independent Spec,CP and Spec,TP in English:**
 a. [TP Stephanie will [be [vP buying the book.]
 b. [CP What will [TP Stephanie [be [vP buying]?
 (3) **Austronesian voice systems:** (from Erlewine, Levin, and Van Urk, to appear)¹
 a. **A privileged argument:** One argument is designated the *subject*, and is realized in a particular morphological form and/or structural position.
 b. **Articulated voice morphology:** Morphology on the verb varies with the choice of subject argument, including options for taking certain oblique arguments as subjects.
 c. **Extraction restriction:** \bar{A} -extraction (*wh*-movement, relativization, etc.) is limited to the subject argument (Keenan and Comrie, 1977, a.o.).
 d. **Marking of non-subject agents:** Non-subject agents are restricted in position or morphologically marked.
 ☞ Voice system lgs seem to conflate T-properties (subjecthood) with C-properties (extractability). We don't observe clearly distinct effects of C and T.

Today:

- I investigate **patterns of preverbal fronting in Toba Batak**, a predicate-initial Austronesian language of northern Sumatra (Indonesia).
- Data here is from elicitation with four speakers currently in Singapore.

Toba Batak follows the Austronesian "subject-only" extraction restriction (Silitonga, 1973; Keenan and Comrie, 1977; Cole and Hermon, 2008, a.o.)

- ☞ I nonetheless argue that **Toba Batak has C and T** with their traditional division of labor (1): C can attract *wh*/focused XPs; T can Case-license and attract DPs.

Patterns of extraction to Spec,CP and Spec,TP are limited, cf English (2):

- (4) **Only non-DPs can move to Spec,CP over Spec,TP:**²
 a. *[CP Aha [TP si Poltak [man-uhor]]]?
 what PN Poltak ACT-buy
 Intended: 'What did Poltak buy?'
 b. ✓[CP Andigan [TP si Poltak [man-uhor buku]]]?
 when PN Poltak ACT-buy book
 'When did Poltak buy a book?'

Idea 1: The asymmetry in (4) is due to the **lack of Case-licensing** for the DP in Spec,CP. (Toba Batak has no case marking, so this is abstract Case.)

What in English (2b) receives accusative low and then moves to Spec,CP.

- (2b) ✓[CP What will [TP Stephanie be buying ACC]?]

But **Toba Batak has no structural Case-licensor in the lower domain of the clause** (e.g. no accusative). I discuss the Case-licensing of in-situ DPs later.

☞ **The limited means of nominal licensing contributes to the observed Austronesian subject-only extraction asymmetry.**

Idea 2: C and T can be **bundled into a single head**. CT probes for targets that are simultaneously *wh*/focused and nominal, and inherits T's Case-licensing ability.

- (5) **CT attracts a *wh*/focused nominal and Case-licenses it:**

[CTP Ise [man-uhor buku]]?
 who ACT-buy book
 'Who bought a book?'

CT also inherits from C the ability to attract multiple targets:

- (6) **CT can front multiple *wh*/focused nominals and Case-license them:**

[CTP Aha [holan si Poltak] [man-uhor]]]?
 what only PN Poltak ACT-buy
 'What did only Poltak buy?'

Note that (6) contrasts minimally from the earlier ungrammatical (4a).

The availability of multiple DP fronting as in (6) has, to my knowledge, never before been documented in any Austronesian voice system language. It is unpredicted by all previous accounts for Toba Batak clause structure (Clark, 1992; Baldrige, 2002; Cole and Hermon, 2008; Hermon, 2009).

☞ **Joint probing by CT derives this privileged status of being simultaneously *wh*/focused and nominal.**

Roadmap §1 Basics §2 Multiple fronting §3 Proposal §4 Spelling out (C)T

²Abbreviations in Toba Batak glosses: ACT(IVE) and PASS(IVE) are voice prefixes. PN precedes proper names. PROX = proximal, MED = medial, DIST = distal.

¹We also observe this constellation of properties in Dinka (Nilotic) (Erlewine, Levin, and Van Urk, 2015, to appear).

1 Toba Batak basics

1.1 Voice and word order

Toba Batak exhibits a two-way voice alternation, similar to nearby Malayic lgs.

- (7) a. **Man**-jaha buku **si Poltak**. (8) a. **Si Poltak** [**man**-jaha buku __].
 ACT-read book PN Poltak PN Poltak ACT-read book
 b. **Di**-jaha si Poltak **buku**. b. **Buku** [**di**-jaha si Poltak __].
 PASS-read PN Poltak book book PASS-read PN Poltak
 ‘Poltak read a book.’ ‘Poltak read a book.’

The **voice prefix** tracks the choice of **subject** (bold). I refer to *maN-* (7a) as **ACTIVE** and *di-* (7b) as **PASSIVE**. Note that the “**PASSIVE**” agent is not demoted or oblique.

Predicate-initial order is the canonical declarative order, but subject-initial order (8) is also common.

☞ Postverbal word order is free with one exception: **the non-subject DP argument must be immediately postverbal** (if any).

(9) Word order with three arguments:

- a. Manga-lehon buku tu si Uli **si Poltak**.
 ACT-give book to PN Uli PN Poltak
 ‘Poltak gave a book to Uli.’
 b. Manga-lehon buku **si Poltak** tu si Uli.
 c. *Manga-lehon tu si Uli [buku **si Poltak** / **si Poltak** buku].
 d. #Manga-lehon si Poltak {tu si Uli buku / buku tu si Uli}.
 ‘The book gave Poltak to Uli.’

(10) Adding *nantoari* ‘yesterday’ to (7a,b):

- a. Man-jaha {*nantoari} buku {nantoari} **si Poltak** {nantoari}.
 ACT-read *yesterday book yesterday PN Poltak yesterday
 b. Di-jaha {*nantoari} si Poltak {nantoari} **buku** {nantoari}.
 ACT-read *yesterday PN Poltak yesterday book yesterday
 ‘Poltak read a book yesterday.’

(11) Agent *wh*-question ⇒ ACTIVE:

- a. ✓ **Ise** [**mang**-allang babi __]?
 who ACT-eat pork
 b. ***Ise** [**di**-allang __ **babi**]?
 who PASS-eat pork
 ‘Who ate pork?’

(12) Theme *wh*-question ⇒ PASSIVE:

- a. ***Aha** [**man**-uhor __ **si P.**?]
 what ACT-buy PN P.
 b. ✓ **Aha** [**di**-tuhor si Poltak __]?
 what PASS-buy PN Poltak
 ‘What did Poltak buy?’

☞ This is the famed “subject-only” Austronesian extraction asymmetry (Silitonga, 1973; Keenan and Comrie, 1977, a.o.).

(13) Extraction of non-DPs does not interact with voice:

- a. ✓ [PP Tu ise] [**man**-uhor buku **si Poltak** __]?
 for who ACT-buy book PN Poltak
 b. ✓ [PP Tu ise] [**di**-tuhor si Poltak **buku** __]?
 for who PASS-buy PN Poltak book
 ‘[For who] did Poltak buy the book?’

(7–13) are my examples but the same patterns have been described by Silitonga (1973), Clark (1984, 1985), Schachter (1984a), and Cole and Hermon (2008).

1.2 *Wh*/focus-fronting

Wh-words prefer to front, but can stay in-situ.³

(14) *Ise* ‘who’

- a. Ise [mang-allang babi on]?
 who ACT-eat pork PROX
 b. Mang-allang babi on ise?
 ACT-eat pork PROX who
 c. Di-allang ise babi on?
 PASS-eat who pork PROX
 ‘Who ate this pork?’

(15) *Andigan* ‘when’

- a. Andigan [man-uhor buku ho]?
 when ACT-buy book you
 b. Man-uhor buku andigan ho?
 ACT-buy book when you
 c. Man-uhor buku ho andigan?
 ACT-buy book you when
 ‘When did you buy the book?’

Passive variants of (15a,b,c) all possible, with positions of *buku* and *ho* reversed.

Phrases with *holan* ‘only’ and *pe* ‘even’ similarly can be in-situ but prefer to move.

(16) *Pe* ‘even’

- a. [Naniura pe] [di-allang ahu].
 naniura even PASS-eat 1sg
 b. Di-allang ahu [naniura pe].
 PASS-eat 1sg naniura even
 c. Mang-allang [naniura pe] ahu.
 ACT-eat naniura even 1sg
 ‘I eat even NANIURA.’⁴

(17) *Holan* ‘only’

- a. [Holan si P.] [mang-allang babi].
 only PN P. ACT-eat pork
 b. Mang-allang babi [holan si P.].
 ACT-eat pork only PN P.
 c. *Di-allang [holan si Poltak] babi.
 PASS-eat only PN Poltak pork
 ‘Only POLTAK ate pork.’

(Why is (17c) ungrammatical but (16c) grammatical? I will return to this.)

NB: At this point, there may be no processes of “*wh*/focus-fronting” *per se*. These fronting examples could simply be due to a general process of optional fronting to preverbal position, with the language *wh*/focus-in-situ at its core.

However, I will show in the next section that *wh* and *holan*/*pe*-focused phrases command a privileged status. I refer to them as “**formally focused**” (**[+FOC]**).

³Both argument and adjunct *wh*-in-situ are not echo questions, as diagnosed by question embedding. See Erlewine (2016) for data.

⁴*Naniura* is special preparation of raw fish pickled in lime juice and spices.

2 Multiple fronting

As we have seen, Toba Batak exhibits the famed “subject-only” Austronesian extraction restriction: if a DP is extracted, it must be the subject (Silitonga, 1973; Keenan and Comrie, 1977; Cole and Hermon, 2008, a.o.).

☞ This predicts that multiple DPs cannot be simultaneously fronted.

[+FOC, +D] [-FOC, +D]:⁵

- (18) *Wh agent, referential DP theme:* (19) *Wh theme, referential DP agent:*
- a. Ise [mang-allang babi ___]? a. Aha [di-tuhor si Poltak ___]?
 who ACT-eat pork what PASS-buy PN Poltak
- b. Babi [di-allang ise ___]? b. Si Poltak [man-uhor aha ___]?
 pork PASS-eat who PN Poltak ACT-buy what
- c. *Ise babi [mang/di-allang ___ ___]? c. *Aha si P. [maN/di-tuhor ___ ___]?
 who pork ACT/PASS-eat what PN P. ACT/PASS-buy
 ‘Who ate pork?’ ‘What did Poltak buy?’

(Cole and Hermon (2008) discuss data such as (18c, 19c) as support for their account.)

[+FOC, +D] [+FOC, +D]: Both can be fronted!⁶

(20) *Wh agent, only theme:*

Ise [holan babi] [*mang/√di]-allang ___ ___]?
 who only pork {*ACT/√PASS}-eat
 ‘Who ate only pork?’ cf (18c)

(21) *Wh theme, only agent:*

Aha [holan si Poltak] [√mang/*di]-allang ___ ___]?
 what only PN Poltak {√ACT/*PASS}-eat
 ‘What did only Poltak eat?’ (=6), cf (19c)

(22) *Multiple fronting of two ‘only’ DPs:*

[Holan posoposo] [holan susu] [*mang/√di]-inum ___ ___].
 only infant only milk {*ACT/√PASS}-drink
 ‘Only infants only drink milk.’

☞ When multiple DPs are fronted, the subject must be immediately preverbal.

⁵The opposite order — [-FOC, +D] [+FOC, +D] — is possible but involves a hanging topic and is not a true multiple extraction. It requires a prosodic break, can be resumed by an optional pronoun (unlike these extractions here), and cannot be embedded. See data in Erlewine (2016).

⁶The order *only < wh* — the opposite of (20,21) — is degraded due to an intervention effect à la Beck (2006). I have thus far not been able to elicit any multiple *wh*-questions.

[+FOC, -D] [±FOC, +D]:⁷

- (23) a. Andigan [(holan) indahan] [*mang/√di]-allang si Poltak ___ ___]?
 when only rice {*ACT/√PASS}-eat PN Poltak
 ‘When did Poltak (only) eat rice?’
- b. Andigan [(holan) si Poltak] [√maN/*di]-tuhor buku ___ ___]?
 when only PN Poltak {√ACT/*PASS}-buy book
 ‘When did (only) Poltak buy the book?’ (=4b)

[+FOC, +D] [-FOC, -D]:

But it’s not simply that any DP and non-DP can be simultaneously fronted:

- (24) a. Ise [man-angko buku [PP sian toko buku] ___ ___]?
 who ACT-steal book from store book
 ‘Who stole books from the book store?’
- b. *Ise [PP sian toko buku] [man-angko buku ___ ___]?
 who from store book ACT-steal book

[+FOC, -D] [+FOC, -D]:

- (25) Boasa [holan [PP sian toko buku]] [man-angko buku ho ___ ___]]?
 why only from store book ACT-steal book 2sg
 ‘Why do you only steal books from the BOOK STORE?’

(26) Summary of multiple extraction patterns:

- a. * [+FOC, +D] [-FOC, +D] V... (18–19)
 b. √ [+FOC, +D] [+FOC, +D] V... (20–22)
 c. √ [+FOC, -D] [±FOC, +D] V... (23)
 d. * [+FOC, +D] [-FOC, -D] V... (24b)
 e. √ [+FOC, -D] [+FOC, -D] V... (25)

☞ The non-subject DP *can* be moved, contra Cole and Hermon (2008) a.o., but only in multiple focused DP extraction (26b).

- Various multiple extractions with non-DPs are possible (26c,e), but still not entirely free (26d).

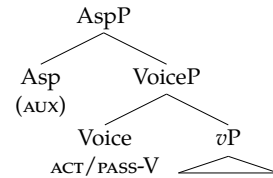
⁷This fronting of the subject to a position between a non-DP *wh* and the verb looks like the Malagasy *bodyguard* construction (Keenan, 1976, a.o.). Ask me about it or see discussion in Erlewine (2016).

3 Proposal

3.1 Clause structure and voice

(27) Working assumptions:

- ACTIVE/PASSIVE is in Voice, above *v*P; the verb head-moves to Voice
- Voice correlates with/determines the **highest DP** in VoiceP:
 - ACTIVE ⇔ agent highest
 - PASSIVE ⇔ patient highest
- Constituents in VoiceP are subject to **scrambling**. All linearizations with the verbal complex leftmost can be generated.



Property (27b) is (roughly) shared with many previous works on Austronesian voice, including Guilfoyle, Hung, and Travis (1992); Aldridge (2004); Rackowski and Richards (2005). Scrambling of postverbal constituents (27c) has been independently proposed for many Austronesian languages, including Malagasy (Paul, 2000; Pearson, 2000), Tagalog (Kroeger, 1991/1993; Richards, 1993; Wegmüller, 1998; Rackowski, 2002), Tongan (Otsuka, 2002, 2005), and Niuean (Clemens, 2014).

3.2 Nominal licensing

Toba Batak lacks case morphology, but **nominals must be licensed** by PF.

- ☞ There is **no structural Case licenser within VoiceP**.
 - The subject is (Case-)licensed by T via Agree.
 - One DP can be **licensed by adjacency**, which allows for postverbal PASSIVE agents and ACTIVE themes.

The idea of licensing by adjacency goes back to Baker (1988, et seq).

(28) Noun incorporation in Southern Tiwa (Allen, 1986, 390):

- Seuanin** i-wan-ban hliawrade-'ay.
man.PL AGR_{3pl}-COME-PAST lady-to
- Am-**seuan**-wan-ban hliawrade.
AGR_{3sg}-man-come-PAST lady
'The men came to the lady.'

Baker (1988, 106ff): 'Come' is unaccusative, so there is only one Case-licensor: T. T Case-licenses with the 'men' in (28a). The 'lady' instead must be an oblique. In (28b), the 'man' incorporates into the verb, freeing T to license the 'lady'.

Massam (2001) observes a similar alternation with NPs:

(29) Pseudo Noun Incorporation in Niuean (ex Clemens, 2014, 90):

- Kua tō he magafaoa e tau huli talo he māla.
PERF plant ERG family ABS PL shoot taro LOC farm
- Kua tō huli talo e magafaoa he māla.
PERF plant shoot taro ABS family LOC farm
'The family planted taro shoots at the farm.'

Massam (2001): Limited to NPs (not DPs): no pronouns or proper names; no possessors, number markers, or indicative relative clauses. Limited to direct objects and instrumentals.

There are also DPs (full nominals) which are licensed by adjacency: non-subject agents in Malagasy and Balinese (see Levin, 2015).

(30) Malagasy non-subject agents must be immediately postverbal; can be proper names, definite (Pearson, 2005, 391ff):

- Vonoin-dRamatoa amin'ny antsy ny akoho.
PASS.kill-Ramatoa with-Det knife Det chicken
'Ramatoa is killing the chickens with the knife.'
- Nohanin' {*haingana} ny gidro {haingana} ny voankazo.
PASS.eat {*quickly} Det lemur {quickly} Det fruit
'The lemur ate the fruit quickly.'

Erlewine, Levin, and Van Urk (2015, to appear): In many Austronesian voice languages, non-subject agents lack structural Case licensing. In some languages, these arguments are licensed through PF adjacency with the verb.⁸

- ☞ **Nominals can be licensed (or waived of the licensing requirement) by PF adjacency with the verb.**

I abstract away from the precise mechanism underlying licensing by adjacency here, but see Levin (2015) for one proposal.⁹ Languages may then vary in the size of adjacency-licensed nominals (N, NP, DP) and possible base positions.

The limited means of nominal licensing explains the word order restrictions observed: The subject will be licensed by T. The non-subject DP will be licensed by adjacency..

(31) Nominal licensing in Toba Batak:

- Man-jaha 🤨_[DP buku] 😊_[DP si Poltak].
ACT-read book PN Poltak
'Poltak read a book.'

⁸In many Formosan and Philippine languages, non-subject agents then receive a last-resort genitive/ergative case. See Erlewine, Levin, and Van Urk (to appear).

⁹In brief, Levin recasts the Case Filter as a requirement that nominals form a part of an extended projection. This could be KP (requiring Case) but postsyntactic adjunction (Local Dislocation; Embick and Noyer, 2001) allows for the nominal to join a verbal extended projection, obviating its need for Case. Postsyntactic adjunction requires head-head adjacency between the highest head of the nominal extended projection and the verb.

We also explain an asymmetry between *holan* ‘only’ and *pe* ‘even’:

(32) **A contrast between *holan*- and *pe*-marked non-subject DPs:**

- a. Mang-allang 🤨_[DP naniura] *pe* 😊_[DP ahu].
 ACT-eat naniura even 1sg
 ‘I eat even NANIURA.’ (=16c)
- b. *Di-allang [holan 🤨<sub>[DP si Poltak]] 😊_[DP indahan].
 PASS-eat only PN Poltak rice
 ‘Only POLTAK ate rice.’ (=17c)</sub>

3.3 C and T

☞ **C and T exist in Toba Batak** with C triggering *wh*/focus-fronting and T associated with subject Case-licensing/fronting.

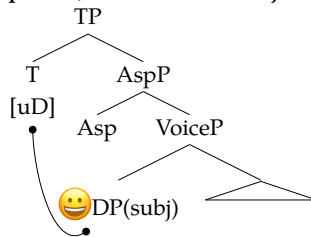
(Assume C and T are unpronounced for now.)

(33) **The content of C and T:**

	C = [uFOC]	T = [uD]
Case-licenses target?	no	yes
Must front target?	yes	no
Can probe multiply?	yes	no

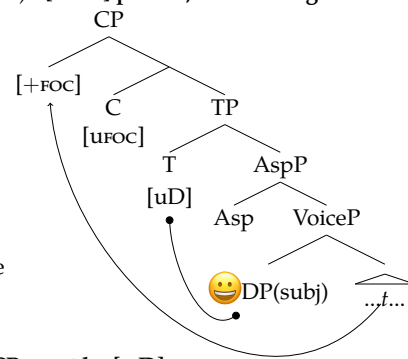
Invocations of these probes are optional.

(34) **[uD] probes, Case-licenses subj:**



The subject in Spec, VoiceP is necessarily the closest DP for [uD] probing from above. [uD] optionally fronts the subject.

(35) **[uFOC] probes, attracts target:¹⁰**



☞ **Anything fronted by [uFOC] to Spec, CP must be [-D].**

C does not Case-license. A non-subject DP could be licensed postverbally by adjacency to the verb, but fronting to Spec, CP would break this adjacency.

This explains the contrast between DPs and non-DPs in Spec, CP (4). ([uFOC] can also probe and attract multiple [+FOC, -D] targets, resulting in pattern (26e).)

¹⁰I presume no Phase Impenetrability effects arising from a VoiceP or vP phase. See discussion in Erlewine (2016).

3.4 Bundled CT

☞ **C and T can be bundled.**

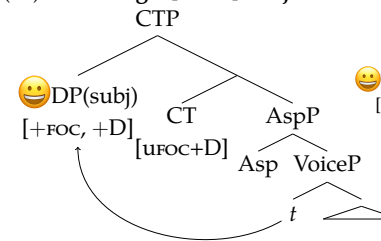
- Bundling of heads accounts for variation in tense, aspect, and mood (Giorgi and Pianesi, 1996), tense and agreement (Bobaljik, 1995; Thráinsson, 1996; Bobaljik and Thráinsson, 1998, a.o.), complementizer systems (Bianchi, 1999), causatives (Pylkkänen, 2002, 2008), V2 requirements (Hsu, 2016, to appear)...
- The CT bundling here is most similar to Legate’s (2011) and Martinović’s (2015) proposals for subject/non-subject extraction asymmetries.¹¹

Their probes get bundled too: **[uFOC+D] probes for targets that are simultaneously [+FOC, +D].**

(36) **A calculus of probe bundling:**

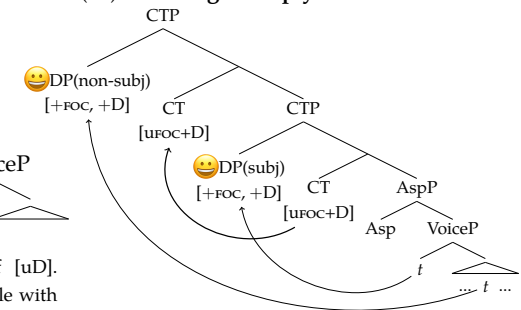
	C = [uFOC]	+ T = [uD]	= CT = [uFOC+D]
Case-licenses target?	no	yes	yes
Must front target?	yes	no	yes
Can probe multiply?	yes	no	yes

(37) **Fronting a [+FOC] subj with CT:**



CT inherits Case-licensing ability of [uD]. (Probing by bundled CT is incompatible with clauses with [-FOC] subjects: the intervening [+D] subject will trigger defective intervention for [uFOC+D].)

(38) **Probing multiply with CT:**



When attracting multiple targets, **CT reprojects.**¹²

Summary:

☞ The traditional organization of C and T — together with a consideration of nominal licensing (abstract Case) and the ability to bundle C and T — derives the patterns of grammatical extractions in the language.

¹¹The Legate (2014) account is couched in terms of *feature inheritance* (Chomsky, 2008); specifically, bundled CT is the result of *under-inheritance*, resulting in traditional T features staying on C. Martinović (2015) instead proposes a dynamic head-splitting process. I do not distinguish between specific implementations of bundling/splitting here.

¹²See Iatridou and Kroch 1992, Watanabe 1992, Browning 1996, and references there on so-called *CP-recursion*. See also more general discussion of head-reprojection in Surányi 2005, Georgi and Müller 2010, and references there.

(39) **Summary, based on (26):**

- a. * [+FOC, +D] [-FOC, +D] V... C doesn't Case-license;
CT would hit defective intervention
- b. ✓ [+FOC, +D] [+FOC, +D] V... multiple attraction by CT
- c. ✓ [+FOC, -D] [±FOC, +D] V... attraction by C and T
- d. * [+FOC, +D] [-FOC, -D] V... C and T are ordered; T can't attract [-D]
- e. ✓ [+FOC, -D] [+FOC, -D] V... multiple attraction by C

We also derive that, when multiple DPs are fronted, the subject must be in immediately preverbal position.

4 Spelling out (C)T

Supporting evidence comes from a ubiquitous, optional particle *na*.

(40) ***na* introducing an embedded clause:**

Hu-boto [(na) modom si Poltak].
 PASS.1sg-know NA sleep PN Poltak
 'I know that Poltak is sleeping.'

Due to examples such as (40), *na* has been called a complementizer (see e.g. Silitonga, 1973). *Na* is also used obligatorily for introducing relative clauses, which I do not discuss here. See Erlewine (2016).

- (41) [+FOC, +D] (✓na) V... (42) [+FOC, +D] (✓na) [+FOC, +D] (✓na) V...
 Ise (✓na) modom? Ise (✓na) holan babi (✓na) di-allang?
 who NA sleep who NA only pork NA PASS-eat
 'Who is sleeping?' 'Who eats only pork?' (na...na ok too)

But there are also cases where my speakers' judgments split, consistently:

- (43) **Pattern A (three speakers):** (44) **Pattern B (one speaker):**
- a. [+FOC, -D] (✓na) V... Andigan (✓na) di-tuhor ho buku-i?
when NA PASS-buy you book-that
'When did you buy that book?'
 - b. [+FOC, -D] (*na) [-FOC, +D] (✓na) V... Andigan (*na) buku i (✓na) di-tuhor ho?
when NA book that NA PASS-buy you
'When did you buy that book?'
 - a. [+FOC, -D] (*na) V... Andigan (*na) di-tuhor ho buku-i?
when NA PASS-buy you book-that
'When did you buy that book?'
 - b. [+FOC, -D] (*na) [-FOC, +D] (*na) V... Andigan (*na) buku i (*na) di-tuhor ho?
when NA book that NA PASS-buy you
'When did you buy that book?'

(45) **Vocabulary insertion rules for optional *na*:**

na or \emptyset \leftrightarrow $\begin{cases} [T] \text{ (also applies to CT)} & \text{(Pattern A)} \\ [C, T] & \text{(Pattern B)} \end{cases}$

The availability of simultaneous *na...na* in (42) supports CT head reprojection.

5 Conclusion

Austronesian voice system languages are an interesting testing ground for our understanding of the left periphery, subjecthood, and extraction:

(3c) **The Austronesian extraction restriction:**

\bar{A} -extraction (*wh*-movement, relativization, etc.) is limited to the subject argument (Keenan and Comrie, 1977, a.o.).

Toba Batak exhibits the Austronesian extraction restriction (3c) (Silitonga, 1973; Keenan and Comrie, 1977, a.o.): If one DP is extracted, it must be the subject, and it generally seems impossible to extract two DPs simultaneously.

☞ **But multiple DPs can be extracted if they are both formally focused.**

The extraction restriction (3c) has been a central topic in the study of Austronesian syntax...and its traditional description might be wrong.

☞ **Nominal licensing** may play a key role in this extraction asymmetry.

- Multiple DPs can be extracted if attracted by the **bundled CT probe**, which is both licensing (from T) and can attract multiply (from C).

Possible multiple extractions of DPs with non-DPs motivate the availability of **split C and T**, with their traditional division of labor; C > T.

☞ **The organization of the clause periphery in Austronesian languages might be much more familiar than previously thought.**

Paper: "Extraction and licensing in Toba Batak" LingBuzz 3261

Handout (with references!): mitcho.com/research/talk-batak-glow.html

This project would not be possible without my Batak teachers, Paris Lubis and Richard Siburian. I thank them for sharing their language with me. I also thank Reinold Limbong and Sopar Amrol Parulian Manik for discussion of judgments. For comments and discussion, I thank Edith Aldridge, Wayan Arka, Louisa Bielig, Mary Dalrymple, Amy Rose Deal, Hadas Kotek, Theodore Levin, Martina Martinović, David Pesetsky, Norvin Richards, Yosuke Sato, Coppe van Urk, Michelle Yuan, and audiences at AFLA 23, the Australian Linguistics Society, Linguistic Society of America, MIT, the University of Hong Kong, the National University of Singapore, and Rutgers. I also thank Nora Samosir, Hannah Choi, and the Choi-Sihombing family for continued encouragement and support of my study of Toba Batak. Errors are mine.

References

- Aldridge, Edith. 2004. Ergativity and word order in Austronesian languages. Doctoral Dissertation, Cornell University.
- Allen, Barbara Jane. 1986. Goal advancement in Southern Tiwa. *International Journal of American Linguistics* 52:388–403.
- Baker, Mark C. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.
- Baldridge, Jason. 2002. Lexically specified derivational control in Combinatory Categorical Grammar. Doctoral Dissertation, University of Edinburgh.
- Beck, Sigrid. 2006. Intervention effects follow from focus interpretation. *Natural Language Semantics* 14:1–56.
- Bianchi, Valentina. 1999. *Consequences of antisymmetry: Headed relative clauses*. Mouton de Gruyter.
- Bobaljik, Jonathan David. 1995. Morphosyntax: The syntax of verbal inflection. Doctoral Dissertation, Massachusetts Institute of Technology.
- Bobaljik, Jonathan David, and Höskuldur Thráinsson. 1998. Two heads aren't always better than one. *Syntax* 1:37–71.
- Browning, Margaret A. 1996. CP recursion and *that-t* effects. *Linguistic Inquiry* 27:237–255.
- Chomsky, Noam. 1986. *Barriers*. MIT Press.
- Chomsky, Noam. 2008. On phases. In *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, ed. Robert Freidin, Carlos P. Otero, and Maria-Luisa Zubizarreta, 133–166. MIT Press.
- Clark, Robin. 1984. The syntactic nature of Logical Form: Evidence from Toba Batak. In Schachter (1984b), 9–16.
- Clark, Robin. 1985. The syntactic nature of Logical Form: Evidence from Toba Batak. *Linguistic Inquiry* 16:663–669.
- Clark, Robin. 1992. Towards a modular theory of coreference. In *Logical structure and linguistic structure*, ed. Cheng-Teh James Huang and Robert Carlen May, 49–78. Kluwer.
- Clemens, Lauren Eby. 2014. Prosodic noun incorporation and verb-initial syntax. Doctoral Dissertation, Harvard.
- Cole, Peter, and Gabriella Hermon. 2008. VP raising in a VOS language. *Syntax* 11:144–197.
- Embick, David, and Rolf Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32:555–595.
- Erlewine, Michael Yoshitaka. 2016. Extraction and licensing in Toba Batak. URL <http://lingbuzz.auf.net/lingbuzz/003261/current.pdf>, manuscript, National University of Singapore.
- Erlewine, Michael Yoshitaka, Theodore Levin, and Coppe van Urk. 2015. What makes a voice system? On the relationship between voice marking and case. In *AFLA 21: The Proceedings of the 21st Meeting of the Austronesian Formal Linguistics Association*, ed. Amber Camp, Yuko Otsuka, Claire Stabile, and Nozomi Tanaka, 51–68. Asia-Pacific Linguistics.
- Erlewine, Michael Yoshitaka, Theodore Levin, and Coppe van Urk. to appear. Ergativity and Austronesian-type voice systems. In *Oxford Handbook of Ergativity*, ed. Jessica Coon, Diane Massam, and Lisa deMena Travis, 373–396. Oxford University Press. URL <http://ling.auf.net/lingbuzz/002629/current.pdf>.
- Georgi, Doreen, and Gereon Müller. 2010. Noun-phrase structure by reprojection. *Syntax* 13:1–36.
- Georgi, Alessandra, and Fabio Pianesi. 1996. *Tense and aspect: From semantics to morphosyntax*. Oxford University Press.
- Guilfoyle, Eithne, Henrietta Hung, and Lisa Travis. 1992. Spec of IP and Spec of VP: Two subjects in Austronesian languages. *Natural Language & Linguistic Theory* 10:375–414.
- Hermon, Gabriella. 2009. Language typology and universal grammar: A commentary on the paper by Eric Potsdam. *Natural Language & Linguistic Theory* 27:773–787.
- Hsu, Brian. 2016. Syntax-prosody interactions in the clausal domain: Head movement and coalescence. Doctoral Dissertation, University of Southern California.
- Hsu, Brian. to appear. Verb second and its deviations: An argument for feature scattering in the left periphery. *Glossa*.
- Iatridou, Sabine, and Anthony Kroch. 1992. The licensing of CP-recursion and its relevance to the Germanic verb-second phenomenon. *Working Papers in Scandinavian Syntax* 50.
- Keenan, Edward. 1976. Remarkable subjects in Malagasy. In *Subject and topic*, ed. Charles N. Li and Sandra A. Thompson. Academic Press.
- Keenan, Edward L., and Bernard Comrie. 1977. Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry* 8:63–99.
- Kroeger, Paul. 1991/1993. Phrase structure and grammatical relations in Tagalog. Doctoral Dissertation, Stanford University.
- Legate, Julie Anne. 2011. Under-inheritance. Presented at NELS 42.
- Legate, Julie Anne. 2014. *Voice and v: Lessons from Acehnese*. MIT Press.
- Levin, Theodore. 2015. Licensing without case. Doctoral Dissertation, Massachusetts Institute of Technology.
- Martinović, Martina. 2015. Feature geometry and head-splitting: Evidence from the morphosyntax of the Wolof clausal periphery. Doctoral Dissertation, University of Chicago.
- Massam, Diane. 2001. Pseudo noun incorporation in Niuean. *Natural Language & Linguistic Theory* 19:153–197.
- Otsuka, Yuko. 2002. VOS in Tongan: passive or scrambling? In *Proceedings of AFLA 9*, 122–136.
- Otsuka, Yuko. 2005. Two derivations of VSO: A comparative study of Niuean and Tongan. In *Verb first: On the syntax of verb initial languages*, ed. Andrew Carnie, Heidi Harley, and Sheila Ann Dooley, 65–90. John Benjamins.
- Paul, Ileana. 2000. Malagasy clause structure. Doctoral Dissertation, McGill University.
- Pearson, Matthew. 2000. Two types of VO languages. In *The derivation of VO and OV*, ed. Peter Svenonius, number 31 in *Linguistic Aktuell*, 327–363. John Benjamins.
- Pearson, Matthew. 2005. The Malagasy subject/topic as an A'-element. *Natural Language & Linguistic Theory* 23:381–457.
- Pylkkänen, Liina. 2002. Introducing arguments. Doctoral Dissertation, Massachusetts Institute of Technology.
- Pylkkänen, Liina. 2008. *Introducing arguments*. MIT Press.
- Rackowski, Andrea. 2002. The structure of Tagalog: Specificity, voice, and the distribution of arguments. Doctoral Dissertation, Massachusetts Institute of Technology.
- Rackowski, Andrea, and Norvin Richards. 2005. Phase edge and extraction: a Tagalog case study. *Linguistic Inquiry* 36:565–599.
- Richards, Norvin. 1993. Tagalog and the typology of scrambling. Honors thesis, Cornell University.
- Schachter, Paul. 1984a. Semantic-role-based syntax in Toba Batak. In Schachter (1984b), 122–149.
- Schachter, Paul, ed. 1984b. *Studies in the structure of Toba Batak*. Number 5 in *UCLA Occasional Papers in Linguistics*.
- Silitonga, Mangasa. 1973. Some rules reordering constituents and their constraints in Batak. Doctoral Dissertation, University of Illinois at Urbana-Champaign.
- Surányi, Balázs. 2005. Head movement and reprojection. *Annales Universitatis Scientiarum Budapestinensis de Rolando Eötvös Nominatae, Sectio Linguistica* 26:313–342.
- Thráinsson, Höskuldur. 1996. On the (non)-universality of functional projections. In *Minimal ideas*, ed. Werner Abraham, Samuel David Epstein, Höskuldur Thráinsson, and Jan-Wouter Zwart, 253–281. John Benjamins.
- Watanabe, Akira. 1992. Wh-in-situ, subjacency, and chain formation. *MIT Occasional Papers in Linguistics* 2.
- Wegmüller, Ursula. 1998. *Sentence structure and ergativity in Tagalog*. Number 36 in *Arbeitspapiere des Instituts für Sprachwissenschaft der Universität Bern*.