

Syntax and Semantics of Japanese Nonconstituent Clefting in Combinatory Categorical Grammar

Yusuke Kubota and E. Allyn Smith
The Ohio State University
kubota@ling.osu.edu, esmith@ling.osu.edu

October 29, 2005

1 Cleft construction in Japanese and nonconstituent clefting

Basic pattern

- (1) Ken ga kono hon o yon-da.
Ken NOM this book ACC read-PAST
'Ken read this book.'
- (2) [Ken ga *t_i* yon-da] no wa [kono hon (o)]_i da.
Ken NOM read-PAST NMLZ TOP this book ACC COP
'It is this book that Ken read.'
- (3) Template: A no wa B da.
A: topic position ('old info')
B: focus position ('new info')

Nominal arguments clefted

Notation: focus position is indicated in *italics* and topic position is indicated in brackets.

- (4) [Ken ga yon-da] no wa *kono hon (o)* da.
Ken NOM read-PAST NMLZ TOP this book ACC COP
'It is this book that Ken read.'
- (5) [Kono hon o yon-da] no wa *Ken (*ga)* da.
this book ACC read-PAST NMLZ TOP Ken NOM COP
'It is Ken that read the book.'
- (6) [Ken ga kono hon o mise-ta] no wa *Ken ni* da.
Ken NOM this book ACC show-PAST NMLZ TOP Mari DAT COP
'It is to Mari that Ken showed this book.'
- (7) [Ken ga kono hon o toriyose-ta] no wa *Huransu kara* da.
Ken NOM this book ACC order-PAST NMLZ TOP France from COP
'It is from France that Ken ordered this book.'

N.B.

- The accusative marker is optional when the object appears in the focus position.

- The nominative marker is obligatorily deleted when the subject appears in the focus position.
- For more oblique postpositions, deletion is impossible.

Multiple argument clefting (Koizumi 2000)

- (8) [Ken ga barasi-te simat-ta] no wa *Mari ni sono himitu o* da.
 Ken NOM disclose EMPH-PAST NMLZ TOP Mari DAT that secret ACC COP
 lit. 'It is to Mari that secret that Ken (inadvertently) disclosed.'
 (which could answer 'What did Ken disclose to whom?')
- (9) [Mari ni barasi-te simat-ta] no wa *Ken ga sono himitu o* da.
 Mari DAT disclose EMPH-PAST NMLZ TOP Ken NOM that secret ACC COP
 lit. ('Who disclosed what to Mari?') 'It is Ken that secret that (inadvertently) disclosed to Mari.'
 (which could answer 'Who disclosed what to Mari?')
- (10) [Sono himitu o barasi-te simat-ta] no wa *Ken ga Mari ni* da.
 that secret ACC disclose EMPH-PAST NMLZ TOP Ken NOM Mari DAT ACC COP
 lit. 'It is Ken to Mari that (inadvertently) disclosed that secret.'
 (which could answer 'Who disclosed that secret to whom?')
- (11) [Barasi-te simat-ta] no wa *Ken ga Mari ni sono himitu o* da.
 disclose EMPH-PAST NMLZ TOP Ken NOM Mari DAT that secret ACC COP
 lit. 'It is Ken that secret to Mari that (inadvertently) disclosed.'
 (which could answer 'Who disclosed what to whom?')

Arguments can be scrambled inside the clefted material.

- (12) [Ken ga barasi-te simat-ta no wa *sono himitu o Mari ni* da.
 Ken NOM disclose EMPH-PAST NMLZ TOP that secret ACC Mari DAT COP
 lit. 'It is that secret to Mari that Ken (inadvertently) disclosed.'
 (which could answer 'What did Ken disclose to whom?')

Combination of argument and adjunct

- (13) [Ken ga Mari ni barasi-ta] no wa *tikasitu de sono himitu o* da.
 Ken NOM Mari DAT disclose-PAST NMLZ TOP basement LOC that secret ACC COP
 lit. 'It is in the basement that secret that Ken disclosed to Mari.'
 (which could answer 'Where did Ken disclose what to Mari?')
- (14) [Mari ni sono himitu o barasi-ta] no wa *Ken ga tikasitu de* da.
 Mari DAT that secret ACC disclose-PAST NMLZ TOP Ken NOM basement LOC COP
 lit. 'It is in the basement Ken that disclosed that secret to Mari.'
 (which could answer 'Where did who disclose that secret to Mari?')

Single adjunct clefting

- (15) [Taroo ga hon o yomi-oe-ta] no wa *tui sakki* da.
 Taro NOM book ACC read-finish-PAST NMLZ TOP just a while ago COP
 'It is just a while ago that Taro finished reading the book.'

Multiple adjunct clefting

- (16) [Taroo ga happyoo-si-ta] no wa *kyonen NELS de* da.
Taro NOM present-do-PAST NMLZ TOP last-year NELS at COP
'It is last year at NELS that Taro presented (a paper).'
- (17) [Taroo ga happyoo-si-ta] no wa *kyonen NELS de Mari to* da.
Taro NOM present-do-PAST NMLZ TOP last-year NELS at Mari with COP
'It is last year at NELS with Mari that Taro presented (a paper).'

Impossible cases

Adjective alone clefted:

- (18) a. Taroo ga *nagai hon o* yon-da.
Taro NOM long book ACC read-PAST
'Taro read a long book.'
- b. * [Taroo ga *t_i hon o* yon-da] no wa *nagai_i* da.
Taro NOM book ACC read-PAST NMLZ TOP long COP
intended: '*It is long that Taro read a book.'

Complex predicate split apart:

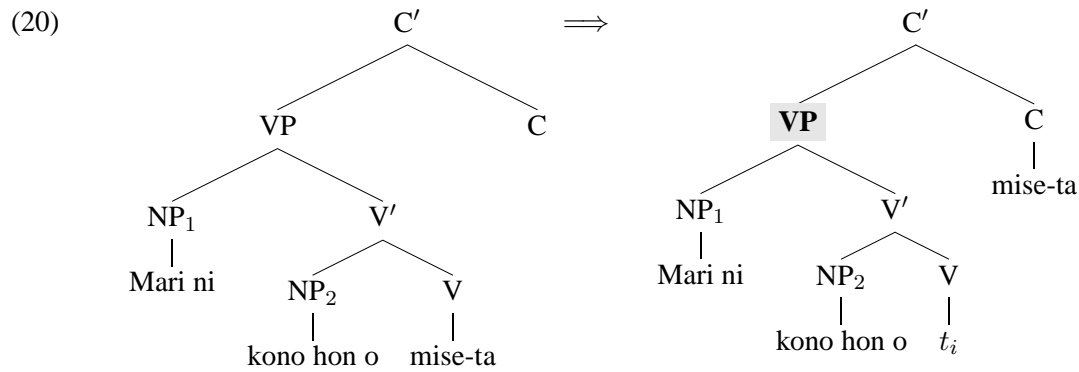
- (19) a. Ken ga Mari ni *sono hon o* yon-de morat-ta.
Ken NOM Mari DAT that book ACC read-MKR BENEf-PAST
'Ken had Mari read that book for him.'
- b. * [Ken ga Mari ni morat-ta] no wa *sono hon o yon-de* da.
Ken NOM Mari DAT BENEf-PAST NMLZ TOP that book ACC read-MKR COP
intended: lit. 'What Ken had Mari do for him was read that book.'
- c. * [Ken ga Mari ni *sono hon o* morat-ta] no wa *yon-de* da.
Ken NOM Mari DAT that book ACC BENEf-PAST NMLZ TOP read-MKR COP
intended: lit. 'What Ken had Mari do for him with that book was read it.'

2 Previous analyses

- Koizumi (2000): verb raising + remnant VP movement
- Takano (2002): oblique movement for nonconstituent formation
- Fukui and Sakai (2003): predicate nominalization ('the other way around')
- Fukushima (2003): floating numerical classifier as the trigger of the nonconstituent

2.1 Koizumi (2000)

- verb raising analysis
- apparent nonconstituent as a VP headed by an empty verb (trace of the raised verb)



(21) Ken ga t_i mise-ta $_j$ no wa [_{VP} Mari ni kono hon o t_j] $_i$ da.
 Ken NOM show-PAST NMLZ TOP Mari DAT this book ACC COP
 lit. 'It is this book to Mari that Ken showed.'

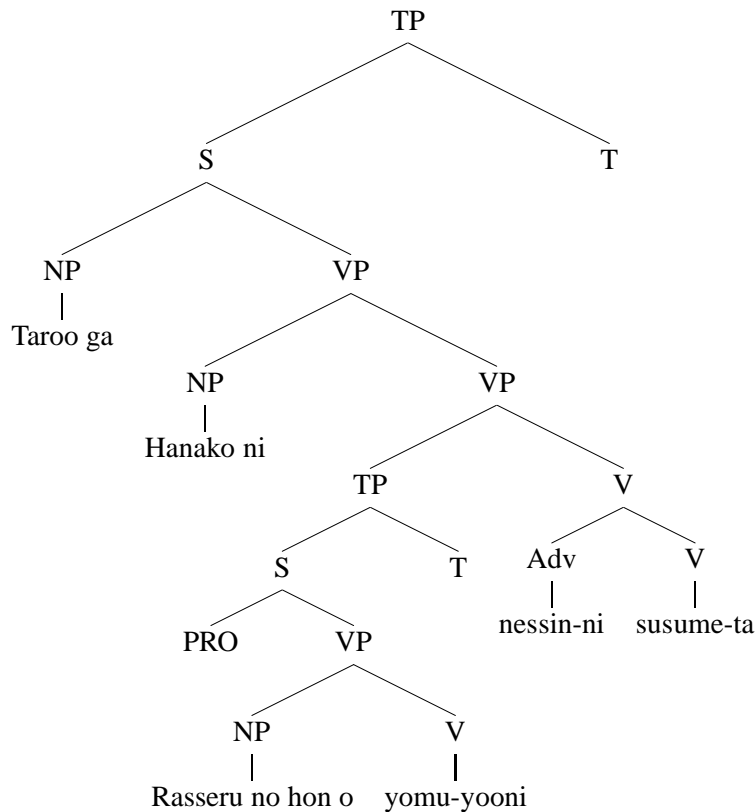
Problems:

- theory-internal inconsistency (see Takano (2002))
- empirical problems pointed out in the literature
 - the 'remnant VPs' don't behave like VPs (Fukushima's (2003) argument regarding the impossibility of VP anaphora as opposed to the possibility of NP anaphora in normal argument clefting)
 - clefting multiple arguments in a certain kind of infinitival complementation construction (Fukui and Sakai 2003).

(22) a. Taroo ga Hanako ni Rasseru no hon o yomu-yooni nessin-ni
 Taro NOM Hanako DAT Russell GEN book ACC read earnestly
 susume-ta.
 recommend-PAST
 'Taro earnestly recommended Hanako to read Russell's books.'

b. [Taroo ga yomu-yooni nessin-ni susume-ta] no wa Hanako ni
 Taro NOM read earnestly recommend-PAST NMLZ TOP Hanako DAT
 Rasseru no hon o da.
 Russell GEN book ACC COP
 lit. 'It is Russell's books to Hanako that Taro earnestly recommended to read.'

(23)



- * The embedded verb has to raise to matrix T for the matrix dative argument and the embedded accusative argument to form a remnant VP, but
- * it couldn't have because the adverb intervenes between the embedded and the matrix verb.

- impossibility of having a completely vacuous VP in the focus position not predicted:

(24) * [Hon o]_k Taroo ga t_i yon-da_j no wa [VP t_k t_j]_i da.
book ACC Taro NOM read-PAST NMLZ TOP COP
intended: lit. 'It is that Taro read the book.'

- impossibility of having an overt VP in the focus position not predicted:

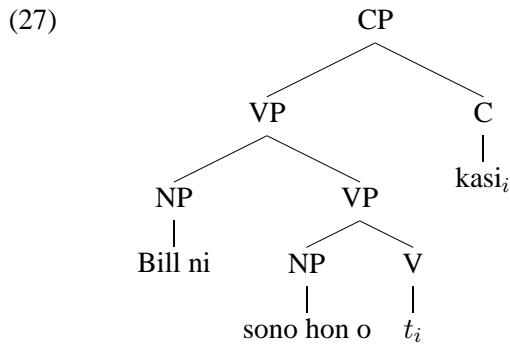
(25) * [Taroo ga t_i] no wa [VP kono hon o yon-da]_i da.
Taro NOM NMLZ TOP this book ACC read-PAST COP
intended: 'What Taro did was read a book.'

This problem can be overcome by

1. making verb raising *obligatory*,
2. assuming that cleft formation applies after verb raising, and
3. stipulating that VPs but not CPs (or IPs) can be clefted.

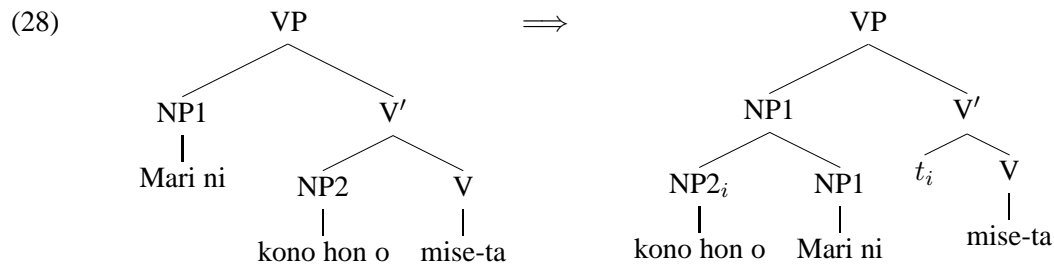
But then another problem arises (because of assumption 1.): how to account for the possibility of coordination of a verbal project containing the direct object and a ditransitive?

- (26) John wa Bill ni sono hon o kasi-te, kono zassi o age-ta.
 John TOP Bill DAT that book ACC lend this magazine ACC give-PAST
 ‘John lent that book and gave this magazine to Bill.’



2.2 Takano (2002)

- oblique movement (adjunction of an NP to another)
- apparent nonconstituent as an NP



Problems:

- Can't account for coordination of nonconstituents easily.

- (29) [John ga ringo o mit-tu] to [Mary ga banana o ni-hon] kat-ta.
 John NOM apple ACC three-CL and Mary NOM banana ACC two-CL buy-PAST
 ‘John bought three apples and Mary bought two bananas.’

Bracketed materials in (29) don't form constituents in the underlying structure.

Takano's proposal: conjoinability of derived constituents.

However,

- this operation is not technically worked out in Takano's (2002) paper (solution not likely to be found)
- once one introduces this operation, not clear how to rule out sentences like the following:

- (30) * Ken wa [NP [NP hon o] sosite [NP Mari ni zassi o]] kat-ta.
 Ken TOP book ACC and Mari DAT magazine ACC buy-PAST
 intended: ‘Ken bought a book (for himself) and bought a magazine for Mari.’

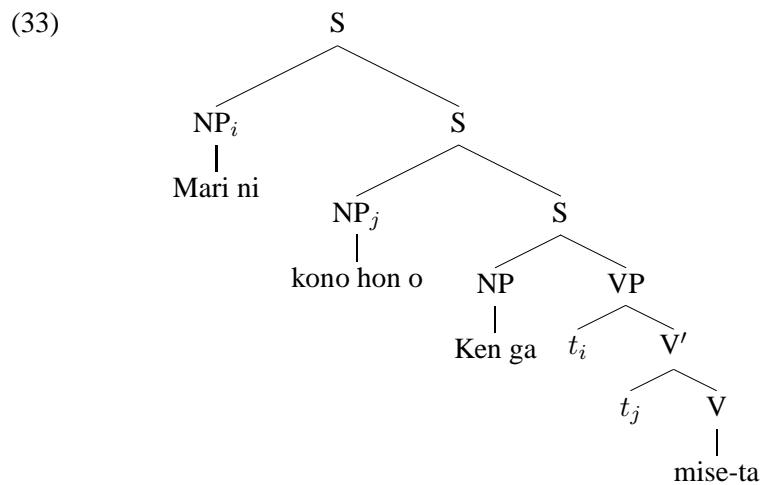
- The ‘bipartite NPs’ don’t really behave like NPs.

- (31) a. Sono hon to kono zassi wa Ken ga Mari ni watasi-ta.
 that book and this magazine TOP Ken NOM Mari DAT give-PAST
 ‘That book and this magazine, Ken gave to Mari.’
- b. * [[Mari ni sono hon] to [Rika ni kono zassi]] wa Ken ga watasi-ta.
 Mari DAT that book and Rika DAT this magazine TOP Ken NOM give-PAST
 intended: lit. ‘To Mari that book and to Rika this magazine, Ken gave.’
- (32) a. Rika ni to Mari ni wa Ken ga kono zassi o watasi-ta.
 Rika DAT and Mari DAT TOP Ken NOM this magazine ACC give-PAST
 ‘To Rika and to Mari, Ken gave this magazine.’
- b. * [Sono hon o Rika ni] to [kono zassi o Mari ni] wa Ken ga
 that book ACC Rika DAT and this magazine ACC Mari DAT TOP Ken NOM
 watasi-ta.
 give-PAST
 intended: lit. ‘That book to Rika and this magazine to Mari, Ken gave.’

2.3 Fukui and Sakai (2003)

Cleft formation in terms of ‘predicate nominalization’:

- scramble (?) the constituents to be clefted to sentence initial position
- then topicalize the rest of the sentence



- c. [Taroo ga aisi-te iru] no wa *zibun no kodomo o* da.
Taro NOM love PRES NMLZ TOP self GEN child ACC COP
'It is his child that Taro loves.' (cleft)
- d. [Zibun no kodomo o aisi-te iru] no wa *Taroo* da.
self GEN child ACC love PRES NMLZ TOP Taroo COP
'It is Taro that loves his child.'
- (38) a. # Hitome mi-ta otoko ga dono zyosi-gakusei ni mo kyuukon-si-ta.
glance look-PAST man NOM which female-student DAT also propose-do-PAST
intended: 'For each female student, the man who glanced at her proposed marriage.'
- b. *Dono zyosi-gakusei ni mo* hitome mi-ta otoko ga kyuukon-si-ta.
which female-student DAT also glance look-PAST man NOM propose-do-PAST
'For each female student, the man who glanced at her proposed marriage.' (scrambling)
- c. # [Hitome mi-ta otoko ga] kyuukon-si-ta no wa *dono zyosi-gakusei*
glance look-PAST man NOM propose-do-PAST NMLZ TOP which female-student
ni mo da.
DAT also COP
intended: 'For each female student, it is the man who glanced at her that proposed marriage.' (cleft)
- The subject can bind the dislocated object in the cleft construction but not in scrambling (37b,c).
 - The bound variable interpretation of the null pronominal is possible in scrambling but not in cleft (38b,c).
 - If dislocation in cleft is mediated by scrambling, the above contrasts between the two constructions is not expected (that is, they should pattern in the same way, but they don't).
- Scrambling and cleft don't exactly parallel with each other in terms of what can be 'displaced'.
Scrambling doesn't obey the Coordinate Structure Constraint (Yatabe 2003):
- (39) a. [Kyoodai to]_i kanozyo ga [Toodai to *t_i*] o kurabe-te iru tte
Kyoto Univ. and she NOM Tokyo Univ. and ACC compare PRES COMP
sit-te ita?
know PAST
'Did you know that she is comparing Kyoto University and Tokyo University?'
- b. [Kyoodai to]_i kanozyo wa [Toodai to *t_i*] o kurabe-te iru no
Kyoto Univ. and she NOM Tokyo Univ. and ACC compare PRES NMLZ
da.
COP
'She is comparing Kyoto University and Tokyo University'

Violation of CSC is not possible in cleft:

- (40) a. * Kanozyo ga [Toodai to *t_i*] kurabe-te iru no wa [*Kyoodai to*]_i
she NOM Tokyo Univ. and compare PRES NMLZ TOP Kyoto Univ. and
da.
COP
intended: '*It is Kyoto University_i that she is comparing *t_i* and Tokyo University.'

- b. Kanozyo ga kurabe-te iru no wa [Kyoodai to Toodai to] da.
 she NOM compare PRES NMLZ TOP Kyoto Univ. and Tokyo Univ. and COP
 ‘It is Kyoto University and Tokyo University that she is comparing.’

2.4 Fukushima (2003)

Lexicalist type-driven analysis (HPSG-like syntax + categorial grammar-like semantics).

Problems:

- Assumes that the numerical quantifier is obligatory in nonconstituent clefting. (Factually wrong.)

- (41) [Ken ga mise-ta] no wa Mari ni hon o (ni-satu) da.
 Ken NOM show-PAST NMLZ TOP Mari DAT book ACC two-CL COP
 lit. ‘It is (two) books to Mari that Ken showed.’

Existence of the numerical quantifier in the focus position *improves* the nonconstituent cleft sentences (Koizumi 2000; Takano 2002), but sentences without them are still acceptable (according to Y.K.’s judgements).

- The analysis proposed by Fukushima (2003) crucially depends on the above assumption.
 - special lexical entry for the numerical quantifier
 - the numerical quantifier glues together the stuff composing the apparent nonconstituent.
- prediction: no numerical quantifier, no nonconstituent
- Doesn’t generalize to the cases without numerical quantifiers.

3 Brief introduction to categorial grammar

For a detailed introduction to the framework of Combinatory Categorial Grammar, see Steedman (1996, 2000b).

3.1 Lexical specification

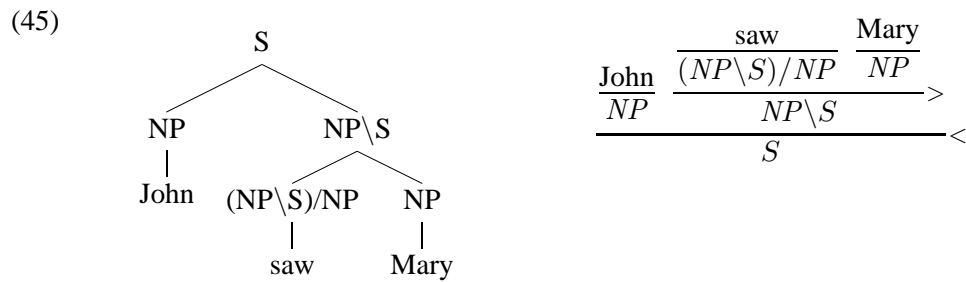
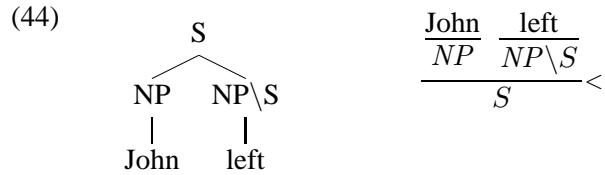
- (42) a. *John*: NP
 b. *Mary*: NP
 c. *left*: $NP \backslash S$
 d. *saw*: $(NP \backslash S) / NP$

- The names of these syntactic categories transparently show what they subcategorize for.
- Lambek-style slash notation: argument under slash (contra Steedman).
- In the absence of parentheses, the forward slash is left associative and the backward slash is right associative.
 E.g.: $NP_a \backslash NP_n \backslash S$ is a shorthand for $NP_a \backslash (NP_n \backslash S)$.

3.2 Logical combinators

Forward/backward application

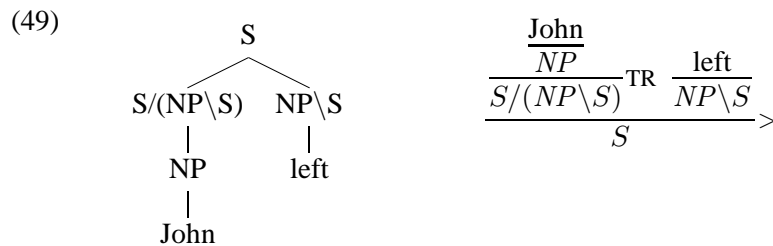
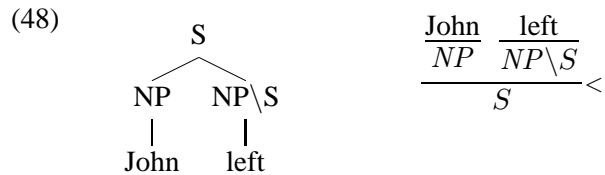
- (43) a. $A/B \ B \vdash A$
 b. $A \ A \backslash B \vdash B$



Type-raising

- (46) a. $A \vdash B / (A \backslash B)$
 b. $A \vdash B \backslash (A / B)$
- (47) $NP \vdash S / (NP \backslash S)$

Example:



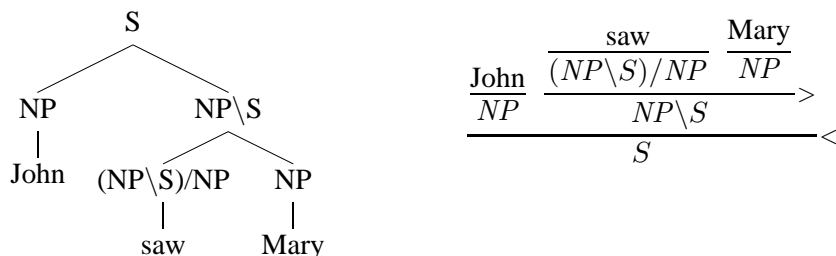
- Reverses the relationship between functor and argument.
- Does not expand the set of sentences admitted by the grammar (except for cases involving certain special constructions such as coordination).

Function composition

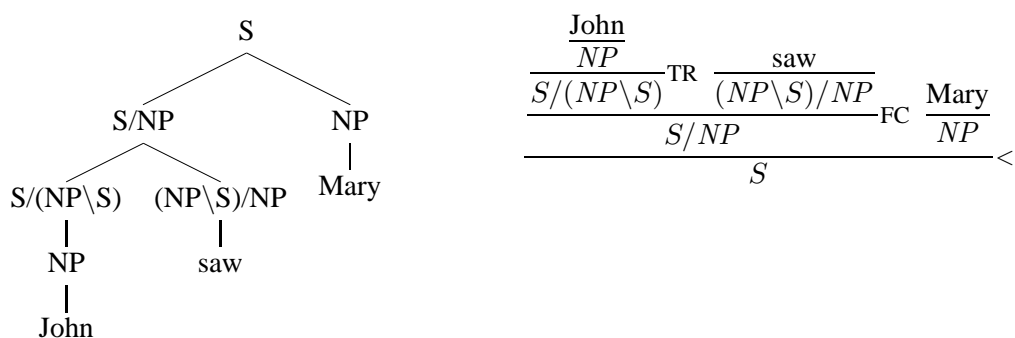
- (50) a. $A/B \ B/C \vdash A/C$
 b. $A \setminus B \ B \setminus C \vdash A \setminus C$

Example:

(51)



(52)



- Note that in the second case, type-raising and function composition yield a surprising constituent.

4 Proposal

4.1 Setting up the stage

- (53) Template: A no wa B da.
 A: topic position
 B: focus position

Lexical entries for function words

- (54) a. $no: (\$ \setminus S_{[-NMLZD, -TOP]}) \setminus (\$ \setminus S_{[+NMLZD, -TOP]}) : \lambda f.f$
 b. $wa: (\$ \setminus S_{[+NMLZD, -TOP]}) \setminus (\$ \setminus S_{[+NMLZD, +TOP]}) : \lambda f.\theta'(f)$
 c. $da: (S_{[+TOP]} / Y) \setminus (Y \setminus S_{[-TOP]}) : \lambda f.\rho'(f)$

Notations:

- \$ is a metavariable ranging over expressions (Ades and Steedman 1982). Two (or more) \$s within an expression stand for the same string of expressions.
 Intuitively: $(\$ \setminus X) \setminus (\$ \setminus Y)$ can be instantiated as
 - $(NP_a \setminus X) \setminus (NP_a \setminus Y)$
 - $(NP_a \setminus NP_n \setminus X) \setminus (NP_a \setminus NP_n \setminus Y)$

- $(NP \backslash PP \backslash X) \backslash (NP \backslash PP \backslash Y)$
- $X \backslash Y$ (where \$ is instantiated as a null string)

but not as

- $(NP_a \backslash NP_d \backslash X) \backslash (NP_a \backslash NP_n \backslash Y)$ (The 2nd argument doesn't match in the case specification.)
- $(NP \backslash PP \backslash X) \backslash (NP \backslash Y)$ (The number of arguments doesn't match.)
- $(NP / X) \backslash (NP \backslash Y)$ (The directionality of the slash doesn't match.)
- Two binary features
 - *NMLZD* (*NOMINALIZED*):
diacritic feature indicating whether a sentence is marked by the nominalizer *no*
 - *TOP* (*TOPICALIZED*):
diacritic feature indicating whether a sentence is marked by the topic marker *wa*
- information-structural markers (Steedman 2000a)
 - θ for theme marker ('old info')
 - ρ for rheme marker ('new info')
- Syntactically almost identity functions with certain constraints and exceptions like the change of feature value/directionality of slash.
- Semantically identity functions (both the function words and information-structural markers).

Lexical entries for verbs

Free word order in terms of lexical ambiguity.

- (55) *tabe-ta*: $NP_a \backslash NP_n \backslash S_{[-N, -T]}$: $\lambda x \lambda y. eat'(y, x)$
tabe-ta: $NP_n \backslash NP_a \backslash S_{[-N, -T]}$: $\lambda x \lambda y. eat'(x, y)$

NP_n (shorthand for $NP_{[CASE\ nom]}$): nominative NP, NP_a : accusative NP, NP_d : dative NP.

4.2 Sample derivations

Single constituent clefting (argument)

- (56) [Ken ga kat-ta] no wa sono hon o da.
 Ken NOM buy-PAST NMLZ TOP that book ACC COP
 'It is that book that Ken bought.'

(57) Syntax:

$$\begin{array}{c}
 \text{Ken ga} \quad \text{kat-ta} \quad \text{no} \quad \text{wa} \quad \text{sono hon o} \quad \text{da} \\
 \hline
 \text{NP}_n \quad \text{NP}_n \backslash \text{NP}_a \backslash S_{[-N, -T]} \quad \frac{(\$ \backslash S_{[-N, -T]}) \backslash (\$ \backslash S_{[+N, -T]}) \backslash (\$ \backslash S_{[+N, +T]})}{(\$ \backslash S_{[-N, -T]}) \backslash (\$ \backslash S_{[+N, +T]})} \text{FC} \quad \frac{\text{NP}_a}{S_{[+T]} / (NP_a \backslash S_{[+T]})} \text{TR} \quad \frac{\text{da}}{(S_{[+T]} / Y) \backslash (Y \backslash S_{[-T]})} \\
 \hline
 \text{NP}_a \backslash S_{[-N, -T]} < \frac{NP_a \backslash S_{[+N, +T]}}{S_{[-T]}} < \frac{(NP_a \backslash S_{[+T]}) \backslash S_{[-T]}}{S_{[-T]}} <
 \end{array}$$

- gapped sentence topicalized (looking for an accusative objection)

- After combining with *no* and *wa*, the gapped sentence is marked as [+N, +T].
- Type-raising on the clefted constitute yields a category that is looking for the gapped S to the *right* to become an S.
- *da* changes th directionality of the slash. That is, after combining with *da*, the clefted constituent is looking for the rest of the sentence to its *left*.
- Another function of *da*: since the final result should be a non-topicalized S, *da* changes the TOP value of the resultant S of the clefted phrase from + to –.

(58) Semantics:

$$\begin{array}{c}
 \frac{\text{Ken ga}}{k} \quad \frac{\text{kat-ta}}{\lambda y \lambda x. \text{bought}'(y, x)} \\
 \hline
 \lambda x. \text{bought}'(k, x)
 \end{array}
 <
 \frac{\frac{\text{no}}{\lambda f. f} \quad \frac{\text{wa}}{\lambda f. \theta'(f)}}{\lambda f. \theta'(f)}_{\text{FC}}
 \frac{\frac{\text{sono hon o}}{\iota x. \text{book}'(x)}}{\lambda P. P(\iota x. \text{book}'(x))}_{\text{TR}}
 \frac{\text{da}}{\lambda f. \rho'(f)}
 <
 \frac{\theta'(\lambda x. \text{bought}'(k, x))}{\rho'(\lambda P. P(\iota x. \text{book}'(x)))}
 <
 \frac{\rho'(\lambda P. P(\iota x. \text{book}'(x)))(\theta'(\lambda x. \text{bought}'(k, x)))}{\text{bought}'(k, \iota x. \text{book}'(x))}$$

- Semantics goes in tandem with syntax. The derivation automatically yields the right compositional semantics by giving uncontroversial denotations to lexical items.
- *wa* identifies the topicalized material as the theme by the θ function.
- *da* identifies the clefted material as the rheme by the ρ function.

(59) Syntax/Semantics together:

$$\begin{array}{c}
 \frac{\text{Ken ga}}{NP_n : k} \quad \frac{\text{kat-ta}}{NP_n \backslash NP_a \backslash S_{[-N, -T]} : \lambda y \lambda x. \text{bought}'(y, x)} \\
 \hline
 NP_a \backslash S_{[-N, -T]} : \lambda x. \text{bought}'(k, x)
 \end{array}
 <
 \frac{\frac{\text{no}}{(\$ \backslash S_{[-N, -T]}) \backslash (\$ \backslash S_{[+N, -T]}) : \lambda f. f} \quad \frac{\text{wa}}{(\$ \backslash S_{[+N, -T]}) \backslash (\$ \backslash S_{[+N, +T]}) : \lambda f. \theta'(f)}}{(\$ \backslash S_{[-N, -T]}) \backslash (\$ \backslash S_{[+N, +T]}) : \lambda f. \theta'(f)}_{\text{FC}}
 \frac{\frac{\text{sono hon o}}{NP_a : \iota x. \text{book}'(x)}}{\lambda P. P(\iota x. \text{book}'(x))}_{\text{TR}}
 \frac{\text{da}}{(S_{[+T]}/Y) \backslash (Y \backslash S_{[-T]}) : \lambda f. \rho'(f)}
 <
 \frac{NP_a \backslash S_{[+T]} \backslash S_{[-T]} : \rho'(\lambda P. P(\iota x. \text{book}'(x)))}{(NP_a \backslash S_{[+T]}) \backslash S_{[-T]} : \rho'(\lambda P. P(\iota x. \text{book}'(x)))}
 <
 \frac{S_{[-T]} : \rho'(\lambda P. P(\iota x. \text{book}'(x)))(\theta'(\lambda x. \text{bought}'(k, x)))}{\text{bought}'(k, \iota x. \text{book}'(x))}$$

Nonconstituent clefting (argument)

- (60) [Ken ga watasi-ta] no wa sono hon o Mari ni da.
 Ken NOM give-PAST NMLZ TOP that book ACC Mari DAT COP
 lit. 'It is this book to Mari that Ken gave.'

Semantics:

$$\frac{\frac{\frac{\text{Ken ga}}{k} \quad \frac{\text{happyoo-si-ta}}{\lambda x.\text{present}'(x)}}{\text{come}'(k)} < \frac{\frac{\frac{\text{no}}{\lambda f.f} \quad \frac{\text{wa}}{\lambda f.\theta'(f)}}{\lambda f.\theta'(f)}\text{-FC}}{\lambda f.\theta'(f)} < \frac{\frac{\frac{\frac{\text{kyonen}}{\text{last-year}'}}{\lambda t.\text{last-year}'(at-nels'(t))}\text{-FC}}{\text{at-nels}'}}{\lambda t.\text{last-year}'(at-nels'(t))}\text{-FC} \quad \frac{\text{NELS de}}{\text{at-nels}'}}{\lambda f.\rho'(f)} < \frac{\text{da}}{\lambda f.\rho'(f)} < \frac{\theta'(\text{present}'(k))}{\rho'(\lambda t.\text{last-year}'(at-nels'(t)))} < \frac{\rho'(\lambda t.\text{last-year}'(at-nels'(t)))(\theta'(\text{present}'(k)))}{\rho'(\lambda t.\text{last-year}'(at-nels'(t)))} < \frac{\rho'(\lambda t.\text{last-year}'(at-nels'(t)))(\theta'(\text{present}'(k)))}{\text{last-year}'(at-nels'(\text{present}'(k)))} <$$

4.3 Resolving cases of overgeneration

The analysis presented here currently overgenerates examples such as (68) – (71).

- (68) * [t_i hon o Taroo ga yon-da] no wa *nagai*_{*i*} da.
 book ACC Taro NOM read-PAST NMLZ TOP long COP
 intended: lit. ‘It is long that Taro read a book.’
- (69) * [Ken ga Mari ni morat-ta] no wa *hon o yon-de* da.
 Ken NOM Mari DAT BENEf-PAST NMLZ TOP book ACC read-MKR COP
 intended: lit. ‘What Ken had Mari do for him was read a book.’
- (70) * John ga Bill ga sono hon o age-ta to omot-te iru no wa *Mary ni*
 John NOM Bill NOM that book ACC give-PAST COMP think PRES NMLZ TOP Mary DAT
 da.
 COP
 ‘It is to Mary that John thinks Bill gave that book.’
- (71) * [Mary ga Bill ni age-ta to it-ta] no wa *John ga sono hon o* da.
 Mary NOM Bill DAT gave COMP say-PAST NMLZ TOP John NOM that book ACC COP
 intended: lit. ‘It is John_{*i*} that book_{*j*} that t_i said Mary gave t_i to Bill.’

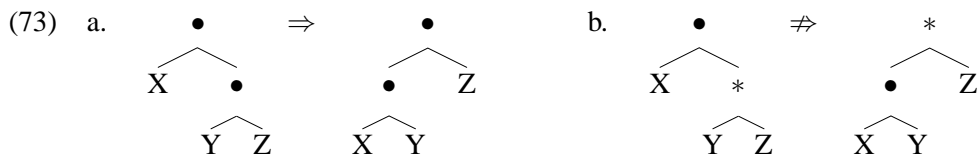
One solution in Multi-Modal CCG (Baldrige 2002)

- Modal control from Type-Logical Grammar (Oehrle 1998) in CCG.
- All language specific constraints on extraction/scrambling in the lexicon.
- Mark islands for syntactic ‘movement’ by a nonassociative mode.

Basics of the solution:

- Two Modes of combination

- (72) a. / and \ for the default associative mode.
 b. /* and * for the nonassociative mode.



- Function composition (relativized to the associative mode)

- (74) a. Forward composition: $A/B \ B/C \vdash A/C$
 b. Backward composition: $A \setminus B \ B \setminus C \vdash A \setminus C$

Crucially, we *don't* have a rule like:

- (75) *Nonassociative forward composition: $A /_* B \ B / C \vdash A / C$

- Fine-tuning combinatoric properties of categories in the lexicon:¹

- (76) a. *nagai*: $NP /_* NP$
 b. *it-ta*: $S \setminus_* NP_n \setminus S$

Returning to examples (68) – (71):

- Illegal derivation of (68) blocked:

(77)

$$\frac{\frac{\text{hon o}}{NP_a} \quad \frac{\text{Ken ga}}{NP_n} \quad \frac{\text{yon-da}}{NP_n \setminus NP_a \setminus S_{[-N, -T]}}}{(NP_a /_* NP_a) \setminus_* NP_a} \quad \frac{NP_a \setminus S_{[-N, -T]}}{NP_a \setminus S_{[-N, -T]}} <$$

*FC

- Illegal derivation of (71) blocked:

(78)

$$\frac{\text{Mary ga Bill ni age-ta to}}{NP_a \setminus S} \quad \frac{\text{it-ta}}{S \setminus_* NP_n \setminus S} *FC$$

- A Multi-modal CCG analysis of the complex predicate construction is under development (based on a work by Kubota and Shin (2005)), which accounts for the ungrammaticality of (69). If you are interested in the details of this, please email us.

¹Motivations for using nonassociative modes for these categories come from the facts that (i) adjectives cannot scramble out of the NP headed by the noun it modifies and (ii) finite embedded clauses resist reassociation of elements inside it with elements in the higher clause:

- (i) *Nagai Taroo ga hon o yon-da.
 long Taro NOM book ACC read-PAST
 intended: 'Taro read a long book.'
- (ii) *Ken ga Rika ga sono hon o, sosite Takasi ga Mari ga kono zassi o yon-da to
 Ken NOM Rika NOM that book ACC and Takashi NOM Mari NOM this magazine ACC read-PAST COMP
 omot-te iru.
 think-MKR PRES
 intended: 'Ken thinks Rika read that book and Takashi thinks Rika read this magazine.'

5 Score Card

How problems for previous analyses are solved / remain to be solved in our analysis:

(√/: solved, *: not solved, ?: not sure at the moment)

Problems	simple version	Multi-modal version
Koizumi (2000)		
the ‘remnant VPs’ don’t behave like VPs	√	√
nonconstituent clefting in infinitival complementation	*	√
vacuous VP in focus position	√	√
overt VP in focus position	√	√
Takano (2002)		
nonconstituent coordination	?	?
topicalization of ‘nonconstituents’	√	√
Fukui and Sakai (2003)		
vacuous VP in focus position	√	√
connectivity effects	*	*
difference in dislocable elements in cleft and scrambling	*	*
Fukushima (2003)		
can’t account for cases without numerical classifiers	√	√

6 Conclusion

- What the Japanese nonconstituent clefting data tell us:
Flexibility with respect to constituency.
- What the CCG analysis tells us:
Lexicon + logic is nearly sufficient.
- The value of multiple approaches to syntax:
Broaches the ‘big issue’.

References

- Ades, Anthony F. and Mark J. Steedman. 1982. On the order of words. *Linguistics and Philosophy* 4(4):517–558.
- Baldrige, Jason. 2002. *Lexically Specified Derivational Control in Combinatory Categorical Grammar*. Ph.D. thesis, University of Edinburgh.
- Dowty, David. 1988. Type raising, functional composition, and non-constituent coordination. In R. T. Oehrle, E. Bach, and D. Wheeler, eds., *Categorical Grammars and Natural Language Structures*, 153–198. Dordrecht: Kluwer.
- Fukui, Naoki and Hiromu Sakai. 2003. The visibility guideline for functional categories: verb raising in Japanese and related issues. *Lingua* 113:321–375.
- Fukushima, Kazuhiko. 2003. Verb-raising and numeral classifiers in Japanese: Incompatible bedfellows. *Journal of East Asian Linguistics* 12:317–347.

- Koizumi, Masatoshi. 2000. String vacuous overt verb raising. *Journal of East Asian Linguistics* 9:227–285.
- Kubota, Yusuke and Hyunsook Shin. 2005. Complex predicates in multimodal categorial grammar: Japanese *-te* form and Korean verbal complex. MS. The Ohio State University.
- Oehrle, Richard T. 1998. Multi-modal type-logical grammar. To appear in R. Borsley and K. Borjars eds., *Non-transformational Syntax*, Blakwell.
- Steedman, Mark. 1996. *Surface Structure and Interpretation*. Cambridge, Massachusetts: The MIT Press.
- Steedman, Mark. 2000a. Information structure and the syntax-phonology interface. *Linguistic Inquiry* 31(4):649–689.
- Steedman, Mark. 2000b. *The Syntactic Process*. Cambridge, Massachusetts: The MIT Press.
- Takano, Yuji. 2002. Surprising constituents. *Journal of East Asian Linguistics* 11:243–301.
- Yatabe, Shūichi. 2003. Does scrambling in Japanese obey the Coordinate Structure Constraint? In *Nihon Gengogakkai Dai-126-kai Taikai Yokōshū*, 262–267. Linguistics Society of Japan.