マルチモーダル範疇文法によるかき混ぜ現象の扱い

小林 昌博 [1]・吉本 啓 [2]

東北大学大学院国際文化研究科

[1] mkoba@insc.tohoku.ac.jp [2] kei@intcul.tohoku.ac.jp

1 はじめに

本論文では，かき混ぜ現象と数量詞遊離現象を範疇文法を拡張した枠組みで説明する方法を提示する．日本語は，語順に対する制約が比較的自由な言語であり，(1) のように語順が入れ替わるかき混ぜ現象が見られる．

(1) a. 太郎が公園で弁当を食べた．

b. 弁当を太郎が公園で食べた．

一方，数量詞遊離現象とは (2) のように名詞句を修飾する数量表現が名詞句から遊離する現象を言う．

(2) a. 学生が3人 酒を飲んだ．

b. 太郎は 本を2冊 買った．

日本語は語順が比較的自由ではあるが，複雑な文や数量詞遊離とかき混ぜの組み合わせなどを含む文では，語順に関する制約が見られる．英語のデータを扱ったかき混ぜ現象の研究として Ross[4] など，日本語に関して Saito[5], [6] などがあるが，数量詞遊離を伴ったかき混ぜ現象を説明できていない．また範疇文法を統語的枠組みとして用いた研究はあまりない．範疇文法は，Montague 意味論 [2] の統語的枠組みとして用いられた経緯があるが，記述力の弱さなどから注目されなかったことがあった。加えて，伝統的な範疇文法では日本語のかき混ぜや数量詞の遊離などを説明することができない．本論文では，範疇文法を拡張した枠組みを用いてかき混ぜ現象と数量詞遊離現象を派生するモデルを提案する．

2 問題の所在

Ross[2] によると，関係節などからなる複合名詞句を含む文などではその名詞句内部の要素が外に出るようなかき混ぜは許されない．

(3) a. John met [$_{NP}$ a child who$_i$ [$_S$ t_i read the book]].

b. * Which book$_j$ did John meet [$_{NP}$ a child who$_i$ [$_S$ t_i read t_j]]?

日本語においても，(1) のような単純な文とは異なり，連体修飾節を含む文では連体修飾節内部から要素を外に取り出すことはできない．

(4) a. 太郎は [$_{NP}$ 典子が秀夫に送った手紙を] 読んだ．

b. * 典子が$_i$ 太郎は [$_{NP}$ t_i 秀夫に送った手紙を] 読んだ．

しかし，(5) に見られるように，連体修飾節内部では要素のかき混ぜが可能である．

(5) 太郎は [$_{NP}$ 秀夫に典子が 送った手紙を] 読んだ．

また，数量詞が遊離した文に関しても，遊離した文にかき混ぜ操作が加わると文法性に違いが出る．

(6) a. 本を 太郎は 4冊 買った．

b. * 学生が 申請書を 3人 提出した．

(6b) と異なり，(6a) はかき混ぜ操作により目的語が文法に出た例であり，目的語と遊離数量詞の間に名詞句が生起できる．範疇文法は，語順の規定が厳密なので，英語などと異なり語順が比較的自由な日本語を記述するには問題がある．実際，(6) や (4b) などの現象に対して，いまだ範疇文法では解決案が出されていない．

3 方法論

本論文では，統語的枠組みとして範疇文法を拡張した枠組みを用いる．伝統的な範疇文法は，言語

	Calculus	*Linguistic Models*
AB	"Classical" (Ajdukiewicz) categorial grammars	binary branching trees
NL	Non-Associative Lambek Calculus	binary branching trees
L	Associative Lambek Calculus	strings
LP	Lambek Calculus with Permutation	multisets

図 2 : categorial systems

表現の統語範疇を関数の組み合わせで表現し，語順を規定する要素が範疇のなかに埋め込まれている．例えば，A と B が範疇ならば A/B と $B\backslash A$ も範疇である．A/B は右手に B の範疇をとり A を返す範疇として働く．$B\backslash A$ はその逆である．この操作を自然演繹表記で表すと図 1 のようになる．派生は線の上から下へと行われる．線の右側にある表記は，その演算で用いられた規則を表す．また，範疇文法では論理学で言う前件肯定規則を「/I, \I」で表す．

$$\frac{A/B \quad B}{A}\,/\mathrm{E} \qquad\qquad \frac{B \quad B\backslash A}{A}\,\backslash\mathrm{E}$$

図 1 ：関数適用

範疇文法の体系は Lambek[1] などにより研究されてきた．図 2 に見られるように，範疇文法の体系は，主に結合則 (*Associativity*) と交換則 (*Commutativity*) あるいは置換 (*Permutation*) の組み合わせにより識別される．生成される言語は，それぞれの体系により異なる．*NL* は結合則と交換則が成立しない体系であり，*L* は結合則は成立するが交換則が成立しない体系である．*LP* は結合則は成り立たないが交換則が成り立つ体系である．範疇文法では，一つの体系のなかで他の体系の「則」を用いてはならない．しかし，これでは複雑な自然言語を記述するには乏しく，特に数量詞の遊離や語順のかき混ぜなど，語彙の移動を伴う現象を記述することはできなかった．例えば，(6) を統語的な観点から見ると，どちらも「名詞句＋名詞句＋遊離数量詞＋動詞」というように同じ語順をとっているが，実際に文法性に違いが生じている．しかし表層的な語順は同じなので，その文法性の違いを表層語順のみで説明することはできない．つまり，表層的な語順のみの組み合わせにより文を派生する伝統的な範疇文法では，(6) の文法性を説明することができない．一方で，交換則と結合則を全体的に認めてしまうと，過剰生成を引き起こすという問題がある．

本論文では，上記の問題点の克服と数量詞遊離と語彙のかき混ぜを派生するためにマルチモーダル範疇文法と呼ばれる Moortgat[3] の枠組みを修正したモデルを用いる．マルチモーダル範疇文法では，$\Diamond$ と $\Box{\downarrow}$ の 2 つのオペレータが NL の体系に加えられる．2 つのオペレータは様相論理に由来し，図 3 のような振る舞いをする．

$$\Diamond \mathrm{A} \vdash \mathrm{B} \Longleftrightarrow \mathrm{A} \vdash \Box{\downarrow}\mathrm{B}$$

$$\frac{\langle \mathrm{A}\rangle \vdash \mathrm{B}}{\mathrm{A} \vdash \Box{\downarrow}\mathrm{B}}\,\Box{\downarrow}\mathrm{I} \qquad \frac{\mathrm{A} \vdash \Box{\downarrow}\mathrm{B}}{\langle \mathrm{A}\rangle \vdash \mathrm{B}}\,\Box{\downarrow}\mathrm{E}$$

図 3 ：オペレータの相互関係

図 3 の上部はシークエント表記によるもので下部は自然演繹表記によるものである．オペレータは，派生の各段階で交換則と結合則の適用を制限する働きをし，過剰生成を抑える役目を果たす．

4 実際の派生

実際にかき混ぜ現象を実現するための語彙規則を設定する．まず，かき混ぜの対象となる名詞句は派生に導入される際に変項として扱われ，カテゴリーに $\Box{\downarrow}$ オペレータが付与される．この操作を以下の必須語彙規則で規定する．α は名詞句の語彙を表す．

$$\alpha \vdash \mathrm{np} \quad \mapsto \quad \mathrm{x} \vdash \Box{\downarrow}\mathrm{np}$$

(6) に見られる，数量詞遊離とかき混ぜの問題を解決するために，以下の環境下でのみ結合則 (*AS*) と交換則 (*MC*) を認める．「$\langle\rangle$」は，$\Diamond$ オペレータが実際の表現にかかった状態を指す．

$$\frac{(\langle \mathrm{A}\rangle \circ \mathrm{B}) \circ \mathrm{C}}{\langle \mathrm{A}\rangle \circ (\mathrm{B} \circ \mathrm{C})}\,\mathrm{AS}$$

$$\frac{\mathrm{A} \circ (\langle \mathrm{B}\rangle \circ \mathrm{C})}{\langle \mathrm{B}\rangle \circ (\mathrm{A} \circ \mathrm{C})}\,\mathrm{MC}$$

デフォルトの語順を「主語＋ (間接目的語) ＋直接目的語＋動詞」と規定する．(6a) の派生は図 4 のようになる．

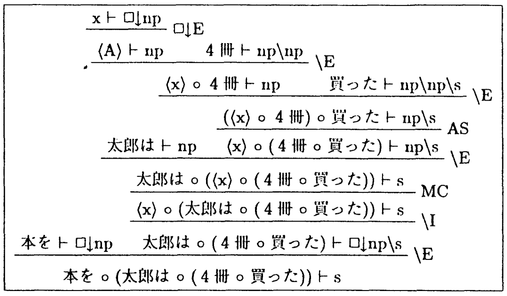

図 4：(6a)「本を太郎は 4 冊買った」の派生

論理学の前件肯定規則に相当する「\, /I」は，交換則が適用された変項のみに適用される.「\, /I」により仮定として派生に導入された変項が消去される．本稿の枠組みにしたがって (6b) を派生すると，上記で規定した交換則と結合則を実行する環境になることはなく，(6b) は適切に排除される．

次に，(4) など連体修飾節を含む例を派生する方法を示す．かき混ぜと同様に移動する名詞句，この場合被修飾名詞句を変項 x で表す．日本語の場合，英語と異なり関係代名詞に相当するものがないので，連体修飾節を作る空の範疇を設ける．

$$\boxed{\epsilon \vdash (\Box{\downarrow}np\backslash s)\backslash np/\Box{\downarrow}np}$$

「ϵ」は左手に動詞句を取り形容詞を返す関数である．(4a) の派生は図 5 のようになる．

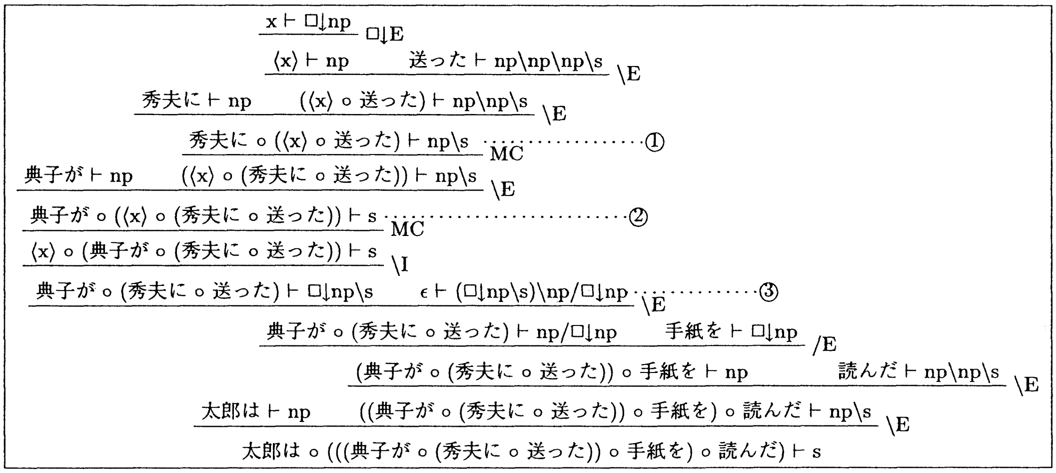

図 5：(4a)「太郎は典子が秀夫に送った手紙を読んだ」の派生

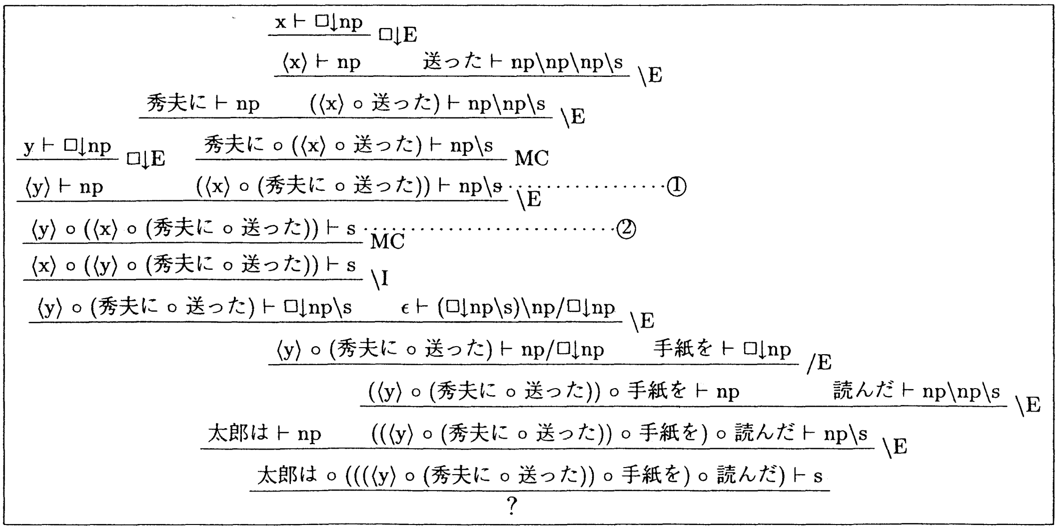

図 6：(4b) *「典子が$_i$ 太郎は t_i 秀夫に送った手紙を読んだ」の派生

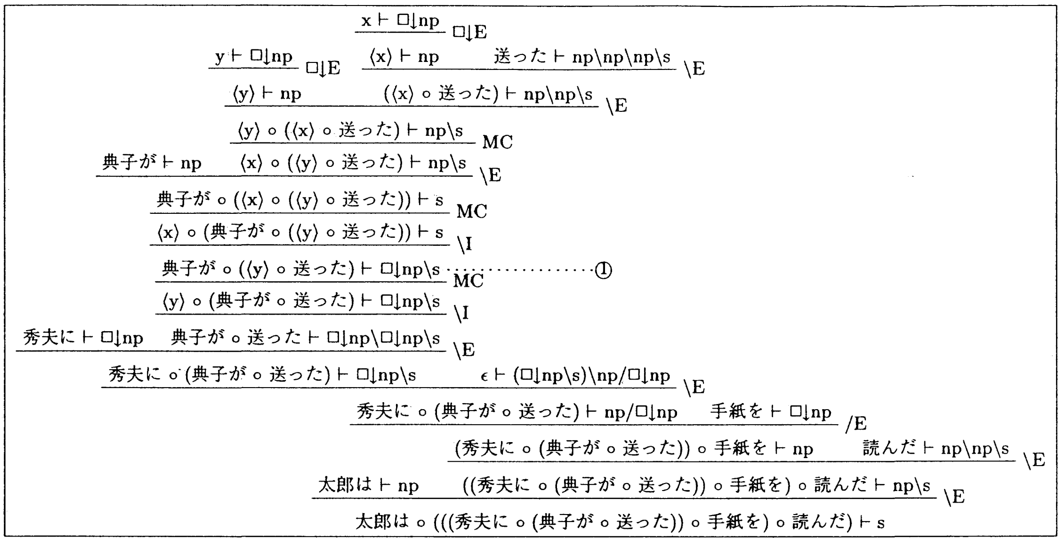

図７：(5)「太郎は秀夫に典子が送った手紙を読んだ」の派生

図５を見ると，①と②の段階で交換則がかかり，被修飾名詞句に相当する x がカッコの外側に出ている．③の段階で引数となる連体修飾節に割り当てられるカテゴリーは，被修飾名詞句が目的語であろうが主語であろうが交換則により常に「$\Box\downarrow np\backslash s$」となる．非文の例である (4b) は，図６であり，名詞句「典子が」が連体修飾節の中から出るパターンなのでそれに相当する仮定 y が用いられる．①と②の段階で交換則が適用されているが，これは仮定 x に対する操作なので，交換則が適用されていない y を消去することができない．結果として y がカッコの中に埋め込まれてしまうので，(4b) のパターンは派生されない．

連体修飾節の内部でかき混ぜが生じている (5) の派生は図７である．(5) は，かき混ぜにより間接目的語「秀夫に」が連体修飾節内部の主語「典子が」の前に出る例である．図７において，①の段階で y に対して交換則が適用されている．したがって，「\I」が適用され y が消去されている．

5 おわりに

本論文では，数量詞遊離現象とかき混ぜ現象を範疇文法を拡張したマルチモーダル範疇文法を用いて派生する方法を提案した．単純な文における数量詞遊離現象とかき混ぜ現象，及び連体修飾節を含む文におけるかき混ぜは，本論文の枠組みで派生することができる．本論文では，埋め込み補文を含む文におけるかき混ぜ現象を扱っていなかったが，今後は扱うデータの拡張と枠組みの修正が必要である．

参考文献

[1] Lambek, J. The mathematics of sentence structure. *American Mathematical Monthly*, Vol. 65, 1958. 154-170.

[2] Montague, R. The proper treatment of quantification in ordinary English. *Approaches to Natural Language*, Reidel, Dordrecht, 1973. 221-42.

[3] Moortgat, M. Constants of grammatical reasoning. *Constraints and Resources in Natural Language Syntax and Semantics*, CSLI, Stanford, 1999.

[4] Ross, J. *Constraints on Variables in Syntax* Ph.D dissertation MIT, 1967

[5] Saito, M. *Some Asymmetrics in Japanese and Their Theoretical Consequence*.Ph.D dissertation MIT, 1985.

[6] Saito, M. Long Distance Scrambling in Japanese. *Journal of East Asian Linguistics*, Vol. 1, 1992. 69-118.