

英語複文の構文解析と編集、その論理と方法

加藤 輝政 小川 清

名古屋市工業研究所
名古屋市熱田区六番町三丁目四番四一号
052-661-3161
tkato@nmiri.city.nagoya.jp

佐良木 昌

佐良木技術翻訳事務所／福岡大学理学部非常勤
名古屋市東区葵二丁目二番二一号
050-150-3755
saraki@st.rim.or.jp

あらまし

英語複文を4パターンに分類した。典型（プロトタイプ）・定型（ステレオタイプ）・従属接続詞を伴わない型（ポテンシャルタイプ）・句形態に従属節構造が潜む型（レイテンタイプ）。この4パターンの複文は、単文結合に還元可能である。分割するのではなく、接続副詞などにより媒介される単文結合に還元する。

キーワード 英語複文、典型、定型、ネクサス、複文の単文結合への還元

Pattern Analyzing of English Complex Sentences and Reducing Hypotaxis to Parataxis

Kato, Terumasa Ogawa, Kiyoshi

Nagoya Municipal Industrial
Research Institute
Nagoya, Atsuta, Rokuban 3-4-41
052-661-3161
tkato@nmiri.city.nagoya.jp

Saraki, Masashi

Saraki Fukuoka University
Office The Department of Science
Part-time Lecturer
Nagoya, Higashi, Aoi 2-2-21
050-150-3755
saraki@st.rim.or.jp

Abstract

English complexes are classified into four main types: prototype, stereotype, potential type, and latent type. The conception of clause is herein interpreted in a broad sense and thus the potential and latent types include NEXUS defined by Otto Jespersen. Hypotaxis which means subordination can be reduced to Parataxis with couplers such as conjunctive adverbs, hence the organization of unconnected coordinated sentences, Parataxis, have been developed to subordination expressing complex relationship between main thought and constituent parts thereof.

key words English complexes, Prototype, Stereotype, Nexus, Hypotaxis, Parataxis

はじめに

英語を母国語としている人たちの、自由で多様な表現や慣用表現を、理論言語学は非文として排除してしまうことがある。慣用として通用している多くの例外文章や破格文は必ずしも非文であるとは言えない。Jespersen 等の伝統文法家達はこの例外的現象を、無視したり切り捨てずに直視して、その例外が生じて来た原因を明らかにしようとした。機械翻訳においては、整合のとれた言語モデルを基準として、生動的で可変的な英語表現を解析しようとする場合、モデルに合わない表現は、解析対象から外れてしまう。そのために、翻訳精度が落ちたり翻訳不能になる。

実際の英文は、いわゆる学校文法や英語5文型の枠には収まらない種々多様な表現形態を伴っている。とくに、二つの事象・観念の相互関係を表現するには、複文が適しているので多用されが、文法通りに、主節と従属節とが従属接続詞を介して接続される構造を成していくとは限らない。たとえば、文法的範疇では、名詞句（たとえば、*the sound of Madonna singing her version Peggy Lee's Fever*）や前置詞句（たとえば、*the season with the remainder getting less*）として扱われる表現形態には、統語構造が潜んでいる。このような表現形態を、離散的なデータとして NP や PP という形に抽象化すると、その統語構造は解析の領域から消えてしまう。

本稿では、複文の諸形態をパターン化し、このパターンを運用することで、複雑な複文を解析する手法を検討する。さらに、従位関係 Hypotaxis を有する英文、すなわち複文は、すべて單文結合 Parataxis と見なすことができるとの観点から、複文の單文結合へ還元する方法について提案する。

I 英語複文のパターン

本稿では、句の形態での表現に、潜在的な統語構造を認ることにより、従属節の範囲を拡張した。この拡張従属節¹には、実詞に後続する分詞節・形容詞節、および形容詞節相当の前置詞句などが含まれる。

英語複文は、單文結合から、種々の中間的段階を経て発展してきた。現代英語における、従属接続詞や関係詞により表される従属構造も、元は、單文同士を、a) 同じ接続副詞を、前出文と後続文とにおいて、反復する形態、b) 後続單文において、先行單文の実詞に相当する代名詞を反復する形態から、発展してきた。L.R.Palmer[1954][1] 小野・中尾[1980][2] 中尾・児島[1990][3] 石橋幸太郎[1972][4] なかでも、ラテン語を規範として、近代英語初期に

¹ 拡張された節という概念、および複文の範疇²については、英語表現学会26回大会における論議の際に、竹内祐二先生（中央大学）から示唆を頂き、これに基づいて設定した。

は、不定代名詞を従属接続詞に転用し、この従属接続詞に媒介された主節・従属節の統合（従属接続詞に導かれる従属節）、不定代名詞を関係詞に転用し、この関係詞に媒介された主節・従属節の統合（関係節）が発達した。これらは複雑な思想を表現しうる高度な表現形態である。

[分析の方法]

英語複文のパターンを、Jespersen [1924][5], Greenbaum, Quirk[1985][6], [1990][7]²及び竹内[1977][8], [1990][9], [1997][10]³などを参考にして、四つのパターン⁴に分類した。そのうち β ・ γ パターンについては、後述の節結合の強度の観点から、下位区分を行った。下位区分 A~D は、複文における節同士の結合度という観点から設定し、その結合度の段階的な強弱レベルを表す（佐良木[1996][11]）。

[複文のプロトタイプ]

従属接続詞が副詞節を導きその先頭に位置する複文を、プロトタイプとし、その節同士の結合強度を標準とした（挿入された従属節は除く）。その結合度を基準として、段階的にレベル分類を行った。

プロトタイプ複文⁵においては、主節が先行し従属節が後続する形態と、従属節が先行し主節が後続する形態との、二形態を探ることができる。このようにプロトタイプ複文では、主節と従属節とが相前後することができる。この程度の主・従節の結合度を、標準とする。この節の移動が可能というのは、プロトタイプ複文の縮約的な形態たる分詞構文や、倒置以外にはみられない（ただし、比較節などの構文には可能なものもある⁶）。一方、以下に述べる β パターンについては、その構文形式は、主節が先行し従属節が後続する形で固定されている。

[複文のパターン]

a 典型複文 (Prototype)

従属接続詞が副詞節を導きその先頭に位置する。

² 節構造を、定動詞節・非定動詞句・無動詞節の三つに分類している。非定動詞節の下位区分は、to つき不定詞、to なし不定詞、-ing 分詞、-ed 分詞としている。

³ 竹内[1981][9]では、名詞に先行する修飾構造を中心に、統語構造を分析している。この論考を手引きに、本稿では、名詞（句）に後続する修飾構造の解析を試みた。

⁴ 言語処理学会第3回年次大会での発表論文では、複文パターンを、統語構造の明示的および暗示的なパターンに分類したが、本稿では再検討し、複文を4パターンに分類した。

⁵ プロトタイプ複文の歴史的概念としての考察は別項を期したい。

⁶ *rather than* の選択節は、先行・後続とも可であるが、*rather than* を複合的な従属接続詞とみなせば、プロトタイプと同じ構文形式である。比較構文などが未検討であるが、今後を期したい。

β 定型複文 (Stereotype)

A. 限定用法の関係節

関係節と主節内の先行詞とは不可分離の統語構造であり、結合度は最強といえる。プロトタイプ複文の従属節のようには、統語的に完全ではない。関係詞が省略されることもある。

B. It that 節

It that 構文は、that 節の一部を that 節の前に切り出した形であり、結合度は標準よりかなり強い¹。

C. 相関節

主節内の先行副詞・形容詞と相関従位接続詞とを介して節同士が相関関係にある so … that ~, such … that ~などの構文を、相関節と仮に呼ぶ。パラタクシスに変換することができるので、結合度は標準よりやや強い。

D. 繼続用法の関係節

主節内の先行詞や文先行詞について、追加的に説明を加えるにすぎ、形態的にもコンマで区切られているので、標準とほぼ同じである。

γ 従属接続詞を伴わない複文 (Potential type)

A. 形容詞用法でかつ後位用法の分詞節

直前に位置する実詞との間で主語－述語の関係があり、形態的には不可分である。しかし、一般に、実詞の一時的状態を記述しているので、限定用法関係節ほど結合は強くない。結合度は αA より弱く αB と同じとする。結合度は標準よりかなり強い。関係節表現の反復を避けるために用いられる（分詞について場合についてのみ言及し、不定詞についての考察は別稿を期したい）。

B. 補部を伴う後位用法の形容詞節

分詞の場合に比して、主語－述語の関係は弱い。また名詞に前位する形容詞に比して名詞との形態的一体化は弱く、一般に実詞の「一時的状態」を記述している。したがって、結合度は、 βA よりも弱い。結合度は標準よりやや強いとする。

C. 比較級反復節 (比例節)

the +比較級…, the +比較級～等の構文をいう。節間の結合を明示する接続詞を欠くものの、文の形式的配列自体が節同士の意味的相関を示す構文である。形態的にはコンマで区切られているので、 βB より結合度は弱い。結合度は標準や強いとする。

D. 分詞構文

接続詞を伴う従属節は主節との時間的または論理的関係を明示するが、このような結節が際立つのを避けた表現である。主節・従属節が前後できるので結合度は標準とほぼ同じとする。

¹ It_wh-節については、未検討である。

δ 句形態に従属節構造を備える複文(Latent type)

名詞+前置詞+NEXUS の形態の名詞句²を含む。たとえば、名詞に後続し、動名詞を目的語とする前置詞句。この句形態は、副詞節として働く独立分詞構文とは異なり、先行名詞を形容する従属節として働き、当該名詞の具体的な性質や付帯的状況を表す。³

[複文の認識論的構造]

英語構文別の認識論的構造は、宮下[12][13]が明らかにしているが、本稿では、この解明に踏まえ、以下の諸点を補足しておきたい。主節と従属節との前後関係、形容詞の後位用法の分析、関係節と非定動詞後位修飾節との併用について、佐良木[1997][14]を参照した。

- ① 典型複文 α では、主節先行・従属節後続が基本で、判断・結論が先述され理由・原因・条件などが後述される表現形式である。逆に、従属節先行・主節後続は、理由原因条件などを先述するので強調表現である。
- ② 定型複文 β は、主節先行・従属節後続の形態に固定されている。概括的抽象的表現からそれを補足する具体的な再表現への叙述過程が基本である。
- ③ 実詞に後続する形容詞節 (γB , δ) とは、実体の概念を先述し（実詞）、その実体の属性・付帯状況・を後述する（形容詞節）表現形態である。
- ④ 同じ文中で、限定用法の関係節 (βA) と非定動詞後位修飾節 (γA) とが、併用される場合には、前者が実詞の恒久的な性質を表し、後者は実詞の一時的性質を表す。
- ⑤ 主節先行・従属節後続は英語散文の基本構造であり、概括的抽象的表現からそれを補足する具体的な再表現という過程が、英語表現の認識論的枠組みであるといえよう。

² 名詞+前置詞に後続する NEXUS には、四種類ある。

α 「派生名詞句」（名詞の動名詞を含む）

β 「形容詞および分詞」+名詞

δ 「動詞的動名詞句」

γ 「名詞+現在分詞（過去分詞、形容詞、名詞、状態前置詞など）」竹内 1997[10]による

³ An ambulance, its engine rumbling loudly, was alongside two unmarked police units with grille lights flashing and three white cruisers with light bars going full tilt.

I went to sleep to the sound of water drumming the slate roof, ... cited from "POSTMORTEN", Patricia D. Cornwell

III 複文編集の方法

[複文の單文への還元]

プロトタイプ複文の、單文結合への還元を、以下のように一般的に表す。
 Clause X, \$Conj Clause Y → SentenceX: \$Adv, Sentence Y.
 \$Conj Clause Y, Clause X → Sentence Y: \$Adv, Sentence X.
 (ここで、\$Conjは任意の従属接続詞、\$Advはこの従属接続詞に意味的に等価の接続副詞である。)

このプロトタイプ複文の還元方式に倣って、 β および γ の各パターンとその結合度とに応じて、そして δ パターンについても、構文形式の意味特性(主節と従属節との結合関係)を表現しうる接続副詞¹や照應関係を明示する語句を補う。この補充により、複文(従属文 Hypotaxis を含む文)を接続副詞・照應語句を介した單文結合 Parataxis へと再編することができる(佐良木・加藤・小川[1997][15])。不完全な構文形式であっても、その構文の統語情報を補いながら原文を單文結合に編集することが可能である。

[接続副詞]

このように、複文を單文結合へ還元する際に、統語情報を付加すると共に、結合度に応じた形態にて單文へと再編する。そのための記述シンボルとして接続副詞を活用する²。接続副詞を活用することで、従属接続詞のもつ意味情報を保持しながら複文を單文結合に還元することが可能となる。一例として、継続用法の関係代名詞節の書換え規則を示す。

<Clause X, \$prep which Clause Y> → <Sentence X. There\$, Sentence Y>

\$prepは任意の前置詞、There\$は接続副詞を表す。この接続副詞は、先行單文と後続單文を意味的に関連付ける語句であり、前置詞+関係代名詞や、関係副詞と意味的に等価である。

There\$:

thereabout, thereafter, thereagainst, thereamong, threat, thereby, therebetween, therefore, threfrom, threin,
 there の付く接続副詞は古語が大半であるが、現代英語でも法律文書や特許文書で盛んに使われている。

[ルール記述の方法]

上記の4パターンを使って、コーパスから、コンコーダンス作成を作成する。作成したコンコーダンスから、当該パターンの適用対象からの除外すべき文例、変形構文例や挿入句を含む文例を発見しながら、①除外対象を

¹ 英語においては、先行文と後続文との結合関係・照應関係を明示する接続副詞が、きわめて多数ある。Kane[1983][5]によれば、18項目百数十個が分類例挙げられる。

² 接続副詞の機能分類を例示する。回帰(return to the point)追加(Addition) 比較と類似(Comparison & Similarity) 讓歩条件(Concession, Qualification) 結論(Conclusion) 結果(Consequence) 反論(Contradiction) 対照(Contrast) 余談(Digression) 例示(Exemplification) 理由(Reason) 反復(Repetition) 連続(Series) 話題転換(Shift of subject) 特定(Specification) 概括(Summation) 時間的継起関係(Temporal relationships) 不確実性(Uncertainty)

記述するルール、②変形・挿入構文などを処理する特殊ルールを作成する。ルールを記述するスクリプトは、常にルールを追加できる構造とする。その上で、一般ルールを、上記パターンを基準に作成する。

一般的ルール構造を示す。

R_{exclusive} ルール適用を除外するルール

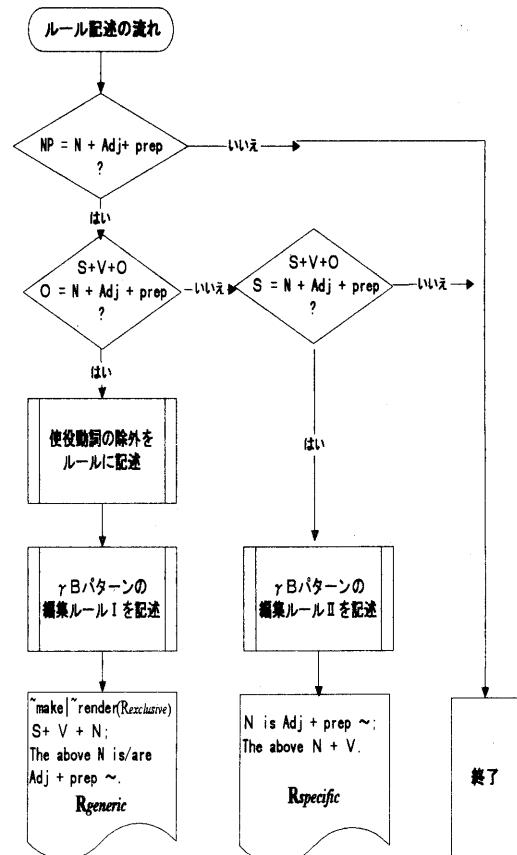
R_{specific} 構文変形についての特殊ルール

R_{generic} 一般ルール

このように、テキスト編集のルール構造は、除外ルール R_{exclusive}、特殊ルール R_{specific}、一般ルール R_{generic} の三層を成す。

C ルール記述の一例

図
ルール記述のプロセス



たとえばγAパターンの一般形態は、動詞の目的語に後続する形容詞節であるが、特殊形態として、主語である名詞に後続する形容詞節がある。一般形態の場合、当該動詞が使役動詞や感情動詞（第5文型を採る）のときは、予め処理対象から外す。

なお、テクスト処理専用言語wraplによるルール記述では、 $R_{exclusive}$ と $R_{specific}$ とを一つのルールに記述できる。例えば、部分的に句構成が同じ構文について、ルール適用を除外するルール $R_{exclusive}$ を $R_{specific}$ に盛り込むことができる。また、wraplの記述では、後位形容詞の直前に、副詞が位置する場合も、γAパターンの一般形態として記述することができる。

以下、図（ルール記述のプロセス）を参照しながらルール記述の具体例について説明する。

R_{generic} の記述

S+V+O : V ≠ Causative Verb, O = N+Adj+ prep ~ の文型のとき、

S+V+O → S+V+N;

The above N is/are Adj + prep ~ という構成でルールIを記述する。wraplでの記述例を次に示す。

1000=(10)+(20 NOT_V CAUSATIVE|~,)+[30 ART] {31 NOUN} [40 ADV]?[41 ADJ_PM] [42 PREP] (43 ~,+/*

10, 11, 20, 30, 31, {;vn}, [The above], 31, [is/are], 40, 41, 42, 43=1001?1001

ここで、波線部分が $R_{exclusive}$ に相当する。

(ここで、ART:冠詞 ADV:副詞 V_CAUSSATIVE:使役動詞 ADJ_PM:後位形容詞 PREP:前置詞)

R_{specific} の記述

S+V+O : S = N + Adj + prep ~, の文型のとき、

S+V+O → N is/are Adj + prep ~; The above N is O.

という構成でルールIIを記述する。wraplでの記述例を次に示す。

1001=[10 ART] (11)*[12 NOUN] [40 ADV]?[41 ADJ_PM] [42 PREP] (43 ~,+/* [50 BE_V|HAVE_V]? (60) +/*

10, 11, 12, [is/are], 40, 41, 42, 43, {;vn}, [The above], 12, 50, 60=e?e

もちろんR_{specific}に記述すべき特殊形態は他にもあり、V = be-Verbではなく一般動詞のとき、さらにその一般動詞の目的語に形容詞節が後続するとき、さらに後位置形容詞の対¹などがある。

¹ 主部内の名詞を修飾する後位置形容詞節で、しかも後位置形容詞の対である例文 To assure the growth of an information infrastructure accessible and accountable to the citizens of the world, governments must develop and implement these objectives in close partnerships with each other and with representatives from business, labor, academia, and the public.

IV 複文編集の実際

以下の例文1・2では、名詞に形容詞節が後続し、かつその形容詞にかかる副詞が、形容詞の直前に位置する。例文3は形容詞が動名詞の補部をとる。これら例文と、IV章で説明したγAパターンの編集ルールに基づいて、単文結合に還元した編集文とについて、機械翻訳による訳出をそれぞれ行った。これら訳文の評価方法としては、

1. 先行名詞と後続形容詞節との統語関係

2. 副詞一後位形容詞との係り受け関係

3. 形容詞補部の解析（形容詞に後方から係る前置詞句）

以上の諸点に関して、「名詞+後位形容詞節」の訳出について、正否を判定することで行った。

例文1・

The key cylinder 32 and bezel 30 may be of substantially conventional designs, which will typically include an actuating portion 36 rotatably movable upon insertion of a key 38 into the key slot 34 and rotation thereof in known lock and key manner.

キー・シリンダー32とベゼル30は、挿入にキー38のキー・スロット34とそれの回転に、周知のロックとキー方法の中で、典型的に回転可能に移動できる動作部分36を含む実質上、従来の設計でありえる。
[評価×₁○₂×₃]

編集文1

The key cylinder 32 and bezel 30 may be substantially conventional designs;
The above will typically include an actuating portion 36;
The above portion 36 is rotatably movable upon insertion of a key 38 into the key slot 34 and rotation thereof in known lock and key manner.

キー・シリンダー32とベゼル30は、実質上、従来の設計でありえる；
上記は、典型的に動作部分36を含む；

上記部分36は、挿入にキー38のキー・スロット34とそれの回転に、周知のロックとキー方法の中で、回転可能に移動できる。
(継続用法の関係節も単文に編集されている)
[評価○₁○₂×₃]

例文2

they subsist on forage on marginal lands generally unsuitable for sustained farming.

訳文

それらは、持続させられた飼育に不適当な一般にぎりぎりの土地の上で、飼料によって生きて行く。

[評価×₁×₂○₃]

編集文2

they subsist on forage on marginal lands;

The above lands are generally unsuitable for sustained farming.

² 機械翻訳には、PC-Transfer/ej（株式会社ノバ）の製品を使った。

それらは、ぎりぎりの土地で飼料によって生きて行く；
上記土地は、一般に持続させられた飼育に不適当である。
[評価○₁○₂○₃]

原文 3

However, these private networks, even the most sophisticated, still suffer from the high cost of leased lines in most countries and the difficulties inherent in attempting to create global networks based on a patchwork of services subject to widely varying capabilities and regulation.

しかし、兵卒がネットワークにのせる、最も多くのものは精巧にしたこれらは、大部分の国の専用回線の高いコストとパッチワークに基づいてサービス主題の広く能力と規制を変化させることにグローバル・ネットワークをつくろうとすることに生まれつきの困難で、まだ苦しむ。

[評価×₁×₂×₃]

編集文 3

However, these private networks, even the most sophisticated, still suffer from the high cost of leased lines in most countries and the difficulties;

The above difficulties are inherent in attempting to create global networks based on a patchwork of services;

The above services are subject to widely varying capabilities and regulation.

しかし、これらの個人のネットワーク（最も洗練されたものでも）は、大部分の国と困難の中で、まだ専用回線の高いコストで苦しむ；

上記困難は、サービスの寄せ集めに基づいてグローバル・ネットワークをつくろうとすることの点で固有である；
上記サービスは、広く能力と規制を変化させることに従う。

[評価○₁○₂○₃]

(Example 1 is cited from cited from U.S. Patent, examples 2 from Oxford MicroConcord Corpus, and example 3 from NII Agenda for cooperation.)

IV 論議

① 主語—述語の統語構造は多様であり、明示的な場合も有れば暗示的な場合もある。NP+VPとして、統語構造を一律に規定すると、統語構造が暗示的な複文の解析は困難である。多様な統語形式を一律に定義すると、具体的な分析は等閑視される。

從属接続詞を伴わない複文γ、潜在的な從属節を備える複文δは、すでにO. Jespersen のNEXUS(対結)定義や、英語学の先駆者が具体的な解明に基づいて、多様な主語—述語の統語形式をナーチンとして把握する試みである。

② 個々の形態素に分解すれば、文のもつ有機的な全体性も失われる。たとえば、補部を伴う後位形容詞を含む文型(以下、NP_AdjP構造という)は、S VOC型という抽象的な文型と同じ形をしている。文をデータとして抽象化して扱う場合、SVOC文型とNP_ADJ_PM構造とは区別が付かない。区別性は失われるわけである。そうであるから、少なくとも、文全体の枠組みを、形態素分解

のまえに、まず把握しておく必要がある。具体的な表現自体を、要素データに分解せずに、表現パターンとして把握することが必要である。池原・宮崎・白井[1996][16]

③ 言語は、社会的諸個人による表現実践の膨大な歴史的集成である。文法的統語や辞書的意味は言語規範であるが、諸個人の自由な表現実践により乗り越えられていく。このように言語は展開的であることをやめないのであるから、自然言語の多様性に対応できる処理機構が必要である。たとえば、本稿で明らかにした複文パターンを運用して、[1]当該パターンの典型的な文型を処理する汎用ルール、[2]同じパターンに属するが、特殊な形や破格の文型をもつ文型を処理するルール、[3]個別処理すべき文型(たとえば、1および2のルールを適用してはならない特定の句や語を排除するため)の処理ルールという三段階の機構を設定することが必要である。

【参考文献】

- [1] L.R.Palmer[1954] "THE LATIN LANGUAGE" Chapter IV, Chapter V
- [2] 小野・中尾 [1980]『英語史I』英語学体系8 大修館書店
- [3] 中尾・児馬 [1990]『歴史的にさぐる現代の英文法』大修館書店
- [4] 石橋幸太郎ら[1972]『現代英語学事典』Hypotaxisの項 成美堂
- [5] Otto Jespersen [1924] "THE PHILOSOPHY OF GRAMMAR"
- [6] Randolph Quirk, Sidney Greenbaum, Geoffrey Leech, Jan Svartvik [1985] "COMPREHENSIVE GRAMMAR OF THE ENGLISH LANGUAGE"
- [7] Sidney Greenbaum, Randolph Quirk[1990] "A STUDENT'S GRAMMAR OF THE ENGLISH LANGUAGE" 池上ら訳『現代英語文法大学編』紀ノ国屋諸点
- [8] 竹内祐治 [1977] 「統語形式の二極性—テニヲカ的統語形式と非テニヲカ的統語形式」英語英米文学第十七集pp.157-178
- [9] 竹内祐治 [1990]「higher wages」などのネクサス性をめぐって」中央大学英米文学会 第7集
- [10] 竹内祐治 [1997] 「"The sound of"などに後続する語群」英語英米文学第36集 pp.157-181中央大学英米文学会
- [11] 佐良木・黒野[1996]「構文の節結合強度に応じた書換え規則によるテキスト処理およびテキスト処理専用言語WRAPLの開発」言語処理学会第2回年次大会発表論文集pp. 161-164
- [12] 宮下眞二 [1982]『英語文法批判』日本翻訳家養成センター
- [13] 宮下眞二 [1985]『英語はどういう言語か』季節社
- [14] 佐良木 昌「自然言語処理からみた英語複文のパターン」英語表現学会第26回大会研究発表ハンドアウト集PP. 19-29
- [15] 佐良木・加藤・小川 [1997] 「WRAPLによる機械翻訳のためのテキスト編集と翻訳精度の向上」言語処理学会第2回年次大会発表論文集pp.
- [16] 池原・宮崎・白井[1997]「言語過程説とALT-J/Eの設計思想」「言語・認識・表現」研究会論文集No.1収録