

# 技術英作文教育に供する対話型日英機械翻訳の設計

野村 芳男 高田 佳悟

東京工業高等専門学校

〒193 東京都八王子市櫛田町 1220-2

**あらまし**：本論文はワープロ感覚で操作する対話型日英機械翻訳に関するものである。本プログラムの具体的な使用は主に高専高学年次教育課程における英作文教育に供することを意図している。英語教育の立場から、本プログラムの基本設計思想は英語文の句による遠心的構造に視点を置き、技術文をこれに対応させて“分かち書き”し、対話形式入力で逐次対訳語句を検索・統合してゆき、翻訳英文を生成するという考え方である。翻訳可能な英文構造、時制・法助詞、日本語文の“分かち書き”入力等に制約条件はあり、辞書も大きくなる傾向がある。しかしながら、訳文品質の良好な技術英文生成が可能である。現在、翻訳できる領域は主に“フリーエ級数”の学術分野である。本論文では、本プログラムの対話型句逐次変換方式による日英機械翻訳の設計思想と翻訳事例を中心に述べている。

**【キーワード】** 日英機械翻訳 技術英作文 CAI 句逐次変換方式 句構造

## On a Program Design of Japanese-English Machine Translation for Making Technical Papers with a sense of Operating English Word Processor

Yoshio NOMURA and Keigo TAKADA

Tokyo National College of Technology

1220-2 Kunugidacho Hachiougi-shi 〒193

**Abstract** The present paper deals with a program of Japanese English machine translation for making technical papers with a sense of operating English word processor. The main objective is to introduce this into a CAI system for the technical English composition at our college. Our developed algorithm is a sequential phrase by phrase translation procedure. In Windows on PC, we can visibly enjoy this procedure from each applied Japanese- phrases to the corresponding English one. In developing this program, we are faced with a lot of difficulties in the translation of the different aspects between Japanese and English. Some of these problems are solved with ideas in the design of Japanese-English translation dictionary. In the present paper, some translated illustrations are disclosed with the design philosophy. Finally, we conclude that this program is helpful to the said CAI.

**key words** Japanese-English machine translation Technical English CAI phrase translation procedure

## 1 まえがき

本論文は英文ワープロ感覚で操作する対話型句逐次変換方式による日英機械翻訳に関するものである。この研究は次の3つの視点から検討している。

①高専高学年次教育課程の技術英作文教育や卒業研究論文の英文概要の作成等に供すること、②英語国人の技術者・研究者に素直に読み理解される英文技術論文や報告書の作成を支援する日英機械翻訳ソフトを設計すること、③対話型句逐次変換方式に関わる技術的課題を抽出すること、④語族の異なる日本語と英語の言語表現の差異から生ずる翻訳限界とその対応策を考究すること等である。ところで、市販日英翻訳ソフトは殆ど一文一括変換方式を適用し、入力文の形態素解析、構文解析および意味解析をすべて完全自動化を施した高度なアルゴリズムを採用したものである。また、多量の日本語文章が高速処理で日英変換される能力を備えている。しかしながら、自然な英語表現の文章化には可成りの手直しが必要であることが現実である。本プログラムの翻訳思想は自然な英語文の生成を目標にし、日英変換後の訳文手直しを少なくすることを狙っている。日英変換アルゴリズムは一文一括変換方式とは異なる。、初等英語の基本構文を修得している学習者（高専生）が文節相当語句を分かち書きしながら順次入力してゆけば、対訳英語語句が検索・統合されて、一文の英語文が生成される句逐次変換方式である。このアルゴリズムは素朴である。対訳辞書は丁寧に設計する必要があるので、大きくなる傾向がある。これまでの研究結果を評価すると、訳文品質は一文一括変換方式と遜色なく、同等あるいは凌駕できる可能性がある。次に、本プログラムを使用した学習者のアンケート調査結果によれば、学習者は①日本語文の求心的構造と英語文の遠心的構造の差異、②英語文の“句構造”による成り立ち、③反復利用による多量の英作文学習が可能になること等を視覚的にも実感でき、英語学習に大変役立つと評価している。技術英作文教育の立場から、対話型句逐次変換方式は技術英作文学習に有効であると考えられる。以下、本論文ではプログラム設計思想と翻訳事例を中心に述べる。

## 2 プログラム設計思想

ここでは、本プログラムの主要な設計事項と目標、それらの基本的な考え方を述べる。

### 2.1 訳文品質

英語国人に素直に読んでもらえる自然な英語表現の訳文品質を目標とする。日本語と英語の言語構造、言語表現、書き手の文章作法、翻訳アルゴリズムの設計限界等に起因する翻訳英文の“良し”、“悪し”は必ずある。これらを認識したうえで、より良質の翻訳英文とすし、手直しを少なくする。

### 2.2 翻訳領域

現段階では“フーリエ級数”の学術分野である。他分野の学術語彙を追加すれば、翻訳可能であるように、拡張性を考慮した設計を目途とする。

### 2.3 翻訳英文構造

単文構造の英文を中心的に設計し、これを複数組み合わせ、並列文、分詞構文、if節、because節、when節、where節等を含む複文も訳出できる設計とする。本プログラムで翻訳可能な英文構造は図1に示す通りである。図1(a)は基本単文構造である。

道標部:「段落」の冒頭、あるいは「文」の文頭に使われる語句を意味する。例えば、“初めに”、“以上纏めると”等である。これらの語句は道路の「交通標識」に相当し、「段落」と「段落」、「文」と「文」とを論理的に繋ぐ。

副道標部:文頭接続詞 [if, since when, where, even if, although...]、慣用的な [it ~ that] 語句、[we ~ that] 語句等は「文」の文意内容の性格を予め示唆する標識語句であるので、本文ではこれを副道標部と呼ぶ。図1(b)は複文の基本構造である。図1(c)は [if A, then B] → [B if A] 等の転置構文である。図1(d)は独立分詞構文等。各部は複数組の名詞句と前置詞句で構成している。これらの組数は英文学術書の調査結果に基づいている[4][5]。図2に各部の(前置詞句+名詞句)組数の調査結果を示す。本プログラムではこのデータを基に

「道標部」= 副詞 または(前置詞句+名詞句)

「提題部」= 名詞句+(前置詞句+名詞句) × 2組

「目的部」= 名詞句+(前置詞句+名詞句) × 3組

の名詞句、前置詞句を訳出する設計としている。

### 2.4 辞書登録する語彙と語句の熟語化

英語国人が書き、定評のある英文学術書文献 [1][2][3]等を利用して、語彙を収集蓄積して辞書登録する。登録語彙は単語レベル、英文法上の慣用句だけでなく、自然な英語表現を訳出するため、出現頻度の高い語句(名詞句、前置詞句、動詞句)を熟語化する[4]。

- (a) **道標部**、**副道標部** **提題部** **動作部** **目的部**、
- (b) **道標部**、**副道標部** **提題部** **動作部** **目的部**、**並列接続詞** **基本単文(a)の繰り返し**  
and, but, 等
- (c) **道標部**、**(パス)** **提題部** **動作部** **目的部** **if, because 等** **基本単文(a)の繰り返し**
- (d) **道標部**、**(パス)** **(パス)** **動作部** **目的部**、**基本単文(a)の繰り返し**

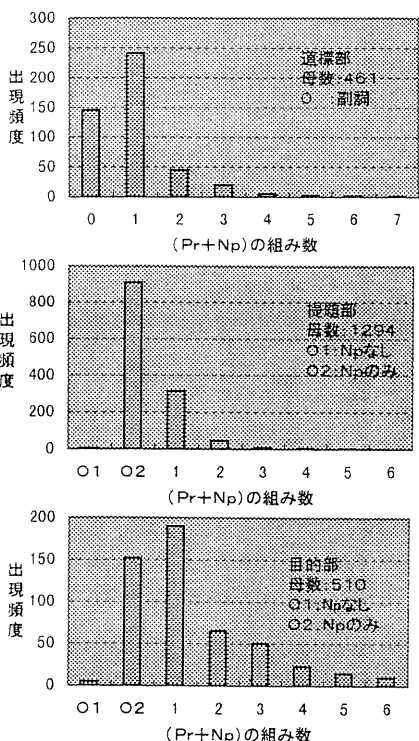


図2 道標部, 提題部, 目的部の(Pr+Np)の分布

## 2.5 翻訳手順の階層化

訳文生成は3段階で“下書き英文”から“推敲英文”の生成を行う。各段階はWindows環境で操作する。第1段階のWindows1の機能(図3):「文」の“下書き”段階と位置ずけて、入力日本語語句の対訳英語語句の検索試行と選定、選定語句の統合化によって「翻訳英文」を生成する。書き手(学習者)が“良し”と判断する翻訳英文(第1次案)を生成するまで、“下書き”を何回も試行できる。第2段階のWindows2の機能(図4):第1次案“下書き英文”を蓄積すると共に、加筆・訂正、推敲を施す作業域と位置ずける。技術英文の“文章”として翻訳英文品質の向上を図る作業域である。

図1 翻訳可能な構文パターン

ここでは、主な作業は①限定詞、単数・複数、法助詞等の検査、②形容詞、副詞等の加筆訂正、③未登録日本語語句の英語化である。第3段階Windows3の機能(訳文事例1):第2段階で蓄積・推敲した訳文をWord97の文書へ転写して、「項」、「節」、「章」単位で英語文章としての推敲、編集等を行う作業域である。更に、この第3段階で「数式」、「定数」、「数式の番号」、「図面、表」等を挿入する。その他、Windows1にはオプション機能(初期設定、文の転置機能、辞書作成機能、登録語句検索機能等)を備える(図5)。例えば、「文」の転置機能は条件文等に施す。具体的には、条件文にはif A, then B. あるいはB if Aの構文がある。このいずれかを初期設定あるいは訳文生成後に選択できるWindowsである。

## 2.6 日本語語句入力のご案内表示

“下書き”作業域のWindows1画面上には、英文基本構造に符合した道標部(文頭修飾語句)、提題部(主部)、目的部、動作部(述部)の順序で、名詞句、前置詞句、動詞句に対応する日本語語句入力を促すを表示する。日本語文の分ち書き(入力語句)は必ずしも文節単位ではなく、準文節相当語句に近い。“案内語”に従って日本語語句を入力すると、対訳変換語句が即座に検索され、日本語語句と対訳英語語句をWindows画面上に逐次的に表示する。対訳英語語句が未登録の場合、日本語語句がそのまま表示される。この操作過程で、学習者は英文ワープロを操作する如く、英語文の句構造を視覚的に捉えることが可能である。Windows1画面上で英作文の下書き感覚で反復的試行により、学習者が“良し”と判断する翻訳英文を完成できる設計とする。次に、“下書き”Windows1における対訳変換手順の一例を示す。例文の“分ち書き”は次の通りである。「初めに、/本章は/フリーエンジニア/の/基礎的なことを/取り扱う。」この順序で日本語入力すると、訳文が生成される。この訳文をWindows2

へ転写し、蓄積する。再び、Windows1 で次の文を分かれ書き、入力して訳文を生成する。このWindowsで、加筆・訂正ができる。数行の訳文を

蓄積すると、Word97 の文書へ転写（訳文事例1）する。Word97 のスペルチェック機能を活用する。

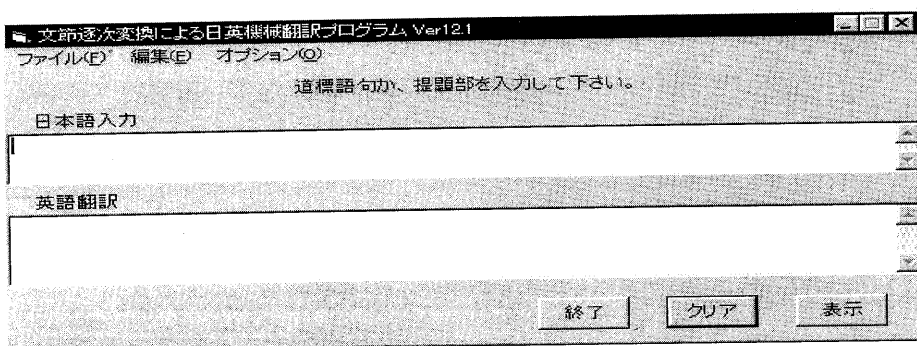


図3 下書き作業域Windows

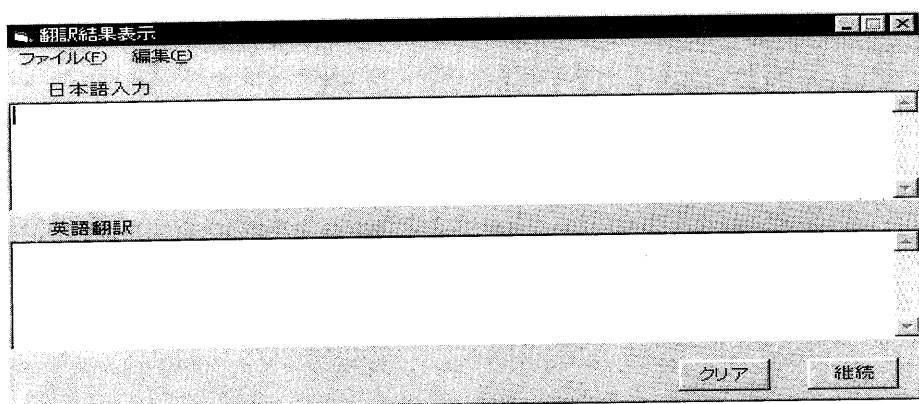


図4 下書き文（1次翻訳文）の蓄積Windows

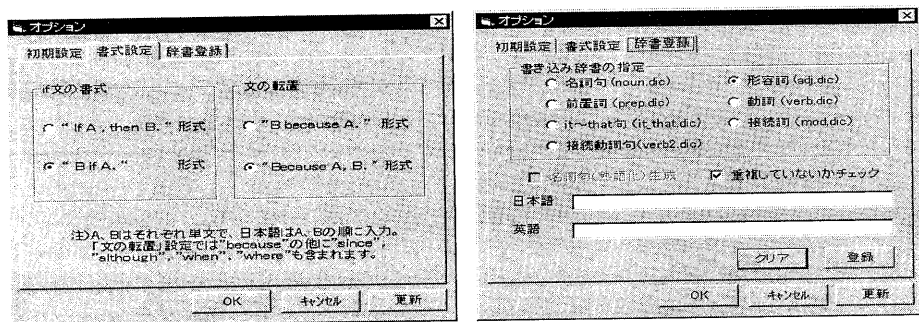


図5 オプション機能Windows

翻訳事例（１） 「文」 の分かち書きと入力操作手順の事例

入力順序

①

道標語句か、提題部を入力して下さい。

日本語入力 “初めに、” を入力する : (G)

②

提題部・前置詞句の目的語が主語を入力して下さい。

日本語入力 “本章は” を入力する : (Np0)

初めに、

英語翻訳

First,

③

前置詞句の目的語を入力して下さい。

日本語入力 “フーリエ級数” を入力する : (Np1)

初めに、本章は

英語翻訳

First, the present chapter.

④

前置詞句を入力して下さい。

日本語入力 “の” を入力する : (Pr1)

初めに、本章はフーリエ級数

英語翻訳

First, the present chapter the Fourier series.

⑤

前置詞句の目的語を入力して下さい。

日本語入力 “基礎的なことを” を入力する : (Np0)

初めに、本章はフーリエ級数の

英語翻訳

First, the present chapter of the Fourier series.

⑥

述部を入力して下さい。単文なら「。」を、複文なら「、」を、続けるならスペースを付加して下さい。

日本語入力 “述べる” を入力する : (V)

初めに、本章はフーリエ級数の基礎的なことを

英語翻訳

First, the present chapter the fundamentals of the Fourier series.

\*\*\*\*\* 英文に翻訳できました! \*\*\*\*\*

訳文完成

日本語入力

初めに、本章はフーリエ級数の基礎的なことを述べる。

英語翻訳

First, the present chapter states the fundamentals of the Fourier series.

## 2.7 未登録語句の検索と辞書登録機能

“分かれ書き”した日本語語句を入力して、対訳語句が未登録の場合、日本語入力語句がそのまま翻訳英文の当該句に出力される。この語句を Windows2 の作業域で英語化もできる。一つの翻訳文は複数の名詞句、前置詞句、動詞句で構成されているから、未登録語句はその言語表現を言い換えて、訳語検索ができる。例えば、「フーリエ級数展開」という名詞句が未登録に遭遇した場合、“下書き” Windows1 で提題部、目的部の名詞句に順次「フーリエ級数」、「級数」、「フーリエ展開」、「展開」と入力すれば、“the Fourier series”, “the series”, “the Fourier expansion”, “the expansion”と訳出される。ここで、新たに、図3の名詞句作成のオプション Windows で、“the Fourier series expansion”を辞書登録して、改めて、当該「文」を翻訳する。あるいは、「フーリエ展開」を選択して翻訳文を生成した後、Windows2 の作業域で“series”加筆する措置を施す。

## 2.8 対訳辞書設計

対訳辞書は道標語句、副道標語句、名詞句(形容詞を含む)、動詞句、前置詞句、形容詞句の各辞書で構成している。各辞書への語彙登録は(a)技術英語文で慣用的に使用される技術用語、(b)文法的な慣用語句、(c)複数単語列が定型化・固定化して用いられる語句等を熟語・慣用語として辞書登録する。現段階では、英語語彙は英文学術書[1][2][3]に現れる語句の統計調査を基にして、出現頻度の高い語句を登録している。辞書設計の具体的な事例は文献[6]に述べている。

## 3 翻訳事例

ここでは、本プログラムで翻訳した幾つかの翻訳事例を示す。翻訳事例2は文献[7]の数式を含む日本語文の翻訳文で、word 文書上で数式を記述している。翻訳事例3は単文、並列文、簡単な関係詞を含む文、条件文、理由文を中心にした日本語文章と翻訳文である。

## 4 学習者の意見・感想

本プログラムを使用した学習者(本校情報工学科5年次 女子学生)の意見感想を述べる。

“技術者のような、論文的文法、専門用語などいつも限られた形の英語文章を書く人にとっては、とて

も便利で、実用的だと思います。特に、その専攻分野に深く進めば進むほど単語は同じものばかり繰り返し用いるようになるし、その度にいちいち辞書を引いてスペルチェックする手間もはぶけると思いますが。論文の場合、日本語でも英語でも、一つの文章が短く簡潔で、分かり易いもの(遠まわしな表現でないもの)がもためられるがちなので、そういった場合は特に有効だと思います。装置の目的について、とても納得しました。もし、輪講などで実用するとしたら、日本語英訳には使えると思います。ただ、辞書にない言葉をどうしても使いたい場合のため、もっと辞書を大きくする必要はあると思います。”

## 5. あとがき

英文ワープロ感覚操作の素朴な句逐次変換方式による技術英作文支援プログラム設計の考え方を述べた。本プログラム設計には幾つかの制約条件を付与しているが、高専教育課程の技術英作文学習に適用できる英語文が生成できる。もとより、英語には、多くの構文、言語表現や語族の差異に基く翻訳限界などが存在する。これらを念頭に置き、本プログラムで翻訳可能な構文の種類、辞書設計、日本語の前処理等を研究し、プログラムの改善を図っている。

## 文献

- [1] A.Papoulis; “Circuit and system – A modern Approach-“Chapter7:FourierSeries pp337-383 Holt Saunders 1981
- [2] A.Oppenheim and R.W.Shafer: “Digital Signal Processing”, Chapter 1:Discrete-time Signals and Systems” pp6-34 Prentice Hall 1985
- [3] S.Soliman “Continuous and Discrete Signals and Systems” Prentice Hall 1993 Chapter 1 Signal Representation pp1-49 Chapter 2 Continuous-Time Systems pp51-116 Chapter 3 Fourier Series pp117-161
- [4] 野村芳男; 英文技術教科書にみる動詞句出現頻度の一視点 信学技報 NLC94-1(1994-05)
- [5] 野村芳男; 前置詞句数からみた英文基本構造の遠心性の一視点 信学技報 NLC95-74 (1996-03)
- [6] 野村芳男、高田佳悟; “数式” 述語「～なる」の自然な英語表現と日英機械翻訳への一視点 信学技報 T L 98-01(1998-05)
- [7] 長瀬道弘、斎藤誠慈; “フーリエ解析へのアプローチ” pp22. 裳華房 1997.10.5 第1版

## 翻訳事例 (2)

### [1] 原文

この関数は偶関数であるから、 $k=1,2,\dots$ に対して

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx dx = 0$$

となる。さらに、

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx = 0$$

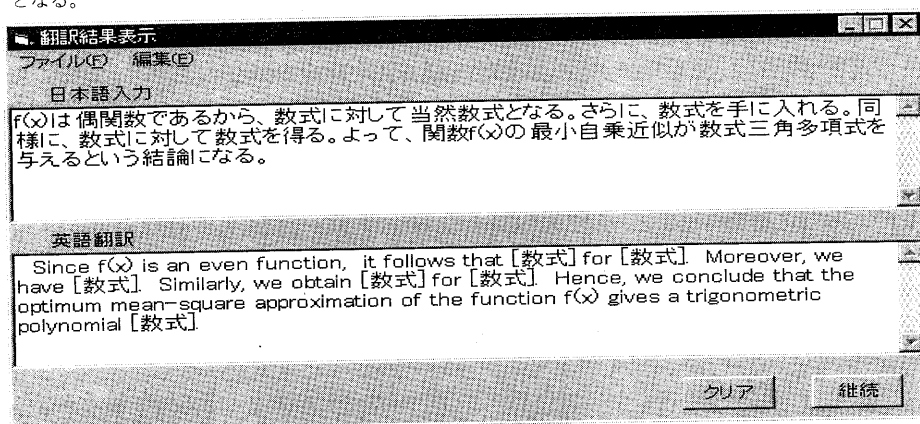
である。また、 $k=1,2,\dots$ に対して

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx dx = \frac{4}{k\pi} \sin \frac{k\pi}{2}$$

よって、 $f(x)$ を最小2乗近似する三角多項式は

$$S_n(x) = \sum_{k=1}^n \frac{4}{k\pi} \sin \frac{k\pi}{2} \cos k\pi$$

となる。



### [3] Word 97 の文書上で具体的な数式記述の事例

Since  $f(x)$  is an even function, it follows that

$$b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx dx = 0$$

for  $k=1,2,\dots$ . Moreover, we have

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx = 0$$

Similarly, we obtain

$$a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos kx dx = \frac{4}{k\pi} \sin \frac{k\pi}{2}$$

for  $k=1,2,\dots$ . Hence, we conclude that the optimum mean-square approximation of the function  $f(x)$  gives a trigonometric polynomial

$$S_n(x) = \sum_{k=1}^n \frac{4}{k\pi} \sin \frac{k\pi}{2} \cos k\pi$$

## 翻訳事例（3）

### (1) 単文

#### 日本語入力

本節における主要課題はいろいろな信号を解析するときに必要な基本道具を明らかにすることである。

#### 英語翻訳

The primary objective in this section is to develop the basic tools required in the analysis of signals.

### (2) 並列文

#### 日本語入力

次例のように、上式は矩形波の余弦級数で、そのフーリエ係数は図におけるように例証される。

#### 英語翻訳

As the following example shows, the above equation is a cosine series of the rectangular waveform, and the corresponding Fourier series coefficients is illustrated as in Fig( ).

### (3) we ~ that 文

#### 日本語入力

前節で、任意周期関数は式で与えられるフーリエ係数の無限和として書かれることを既に示した。

#### 英語翻訳

In the preceding section, we have shown that an arbitrary periodic function can be written as an infinite sum of the Fourier series coefficients given by Eq( ).

### (4) 簡単な関係詞を含む文

#### 日本語入力

信号解析の中核を形成するフーリエ理論はあらゆる工学分野において基本的に重要なものである。

#### 英語翻訳

The Fourier theory that forms the core of the signal analysis is of fundamental importance in all engineering disciplines.

### (5) 条件文

#### 日本語入力

$f(t)$ が一実関数であるならば当然、そのフーリエ係数は実である。

#### 英語翻訳

It follows that if  $f(t)$  is a real function, the corresponding Fourier coefficients are real.

### (6) 理由文

#### 日本語入力

上式の右辺における積分が零になるため、これは次式を生む。

#### 英語翻訳

This yields the following equation because the integral in the right hand side of the above equation tends to zero.