

英語英文学・文献学とコンピュータ
『妖精の女王』1590年版の植字工同定を例に

佐藤治夫
日本大学歯学部

英語英文学、文献学上のテキストデータ処理にあたり、この分野固有の問題の解決が必要となる。16世紀末の英文学史上の金字塔『妖精の女王』1590年版テキストの植字工同定研究過程を例として紹介して、現代語以外のテキストデータの使用に当たったの諸問題を検討し、併せて英語英文学テキストデータの利用方法を探る。

APPLICATION OF COMPUTERS IN THE FIELDS OF ENGLISH LITERATURE
AND TEXTUAL CRITICISM
--With Special Reference to the Case of Compositor Detemination in
The Faerie Queene (publ. 1590)--

Haruo Sato

Nihon University School of Dentistry

1-8-13 Kanda-Surugadai Chiyoda-ku Tokyo 101 JAPAN

As an example of computer-assisted scholarship in the fields of English literature, linguistics and textual criticism, a compositor-determination study on The Faerie Queene (publ. 1590) is described. The author presents some of the difficulties he encountered during the text-input process and its analysis. He finishes by evaluating the situation of computerization in the same fields.

【まえがき】

我国の英語英文学（英語学と英文学）に於ける電算機の本格的な利用の歴史はかなり新しく、現在普及段階である。電算機を使用した初めてのコンコーダンス（用語総索引）は、ハーバード大学によるシェイクスピアコンコーダンスであったが、使用したテキストデータは学外の研究者には提供されず、単にコンコーダンスの出版そのものの為に入力した結果となっている。日本国内の研究者は、お互いの連携が取れていないことから、各研究者が入力したテキストデータを個人レベルでしか利用できない現状にある。

現時点では、どの研究者が何という作品を入力したか不明なままといった状態である。しかし、1984年に「計算機利用言語学研究会」が研究者対象に行なったアンケート形式による調査によって、一部ではあるが電算機メディアに記録された文学（各国文学を含めて）テキストの所在が明らかにされている。〔1〕

調査が行なわれた後に、現状のままでは同じテキストを別々の研究者が互いに連絡の取れない内に入力をしてしまう可能性を指摘した文書を調査に回答した研究者に送り、何がしかの調整をおこなうべく研究会を設立して今日に至っている。現在の会員数は十余名であり、その全てが数メガバイトのテキストデータを保有している。しかしながら、現段階ではテキストデータの利用方法は主にコンコーダンスなどでの利用にとどまっている場合が多いのは残念である。本発表ではテキストデータ処理の一つとして、英文学史上の大作エドモンド・スペンサー作『妖精の女王』1590年版の植字工同定作業の一部を紹介し、電算機を英語英文学研究に利用する場合の諸問題を検討する。

【研究概要】

本研究は『妖精の女王』の1590年版（初版）の活字を拾った植字工の分担ページの同定を目指すものである。当時の出版を考えると無理からぬことであるが、活字になった書物も、校正作業が十分でないために、必ずしも著者の意図を反映

した文章、単語が書いてあるとは限らないところが研究者泣かせである。現在の英語のように正字法(orthography)が確立される以前であったので、活字を拾った職人も自分の好きなように綴っていた部分も入っており、いわば誤植と勝手に変えた綴りの単語が入り交じった部分があり、作者の意図が正確に伝わっているかどうかとも疑わしい。このため、電算機にテキストを入力して、特徴ある綴り方のページ毎の出現パターンを調査したものである。

《近代前期英語のテキストについて》

正字法の確立する以前には、どの言語も原則として発音通りに記述するのは当然であった。特に英語では、ギリシャ語やラテン語のような格変化語尾が消滅する過程がその重要な点であるので、語尾が一定しない。そのために、使用する文字種は殆ど現代の英語と変わらないが、'charitie'のように語尾の綴りが現代と異なるものが多く含まれている。また、ラテン語の影響でJ, Uは、それぞれI, Vで表記されることが多い。

著作権法が施行される以前のことであるので、作品の出版には著者が介在しない場合もあり、校正作業は印刷所の親方まかせといった例も見られる為にテキストそのものの信頼性が疑問視されることもある。

《テキストの入力》

『妖精の女王』は16世紀の長詩であり約1万5千行のテキストの分量である。この作品を入力するにあたり最も研究者を悩ませたのが、ASCIIコードに無い文字の存在であった。近代初期英語までは、ドイツ語に見られるロングエスが存在する。これは英語英文学の研究者のみならずASCIIコードに存在しない文字を持つ言語とその文学作品の研究者に共通の悩みである。一般には、他の文字（@など）で置き換えれば済むと言われるが、i l l n e @ @ e s (illnesses) などという情けない文字列を見せつけられた、「これから研究に電算機を利用しようとする英語英文学研究者」

の反応は予想するまでもない。また、この作品のような韻文を、活字を手で拾って印刷する場合に1行に詰め込もうとして、特殊な記号を使用する例もある。フォントの自由度の向上が望まれる点である。

本研究では、タイピストを雇用して入力作業を行なわせたが、オリエンテーションの為にタイピストが交代する度に2日程度削いてロングエスはエスに、'vp' 'ivst' など、現在の英語とは異なる部分をそのまま入力してもらおうべく教育を行なった。テキストデータの入力が終了すると校正作業が開始され、さらに半年程度の時間が経過した。校正作業の時間短縮のためにテキストファイルをスペルチェッカーにかけてみたが余りにミスペリングとしてはねられる単語が多く実用とはならなかった。スペルチェッカーは正字法が確立した後の時代(それ以前は原則として発音通りに綴っていた)の文献にのみ有効なことを改めて確認した。[2]

《テキストの処理》

出現パターンはページ毎であるがバジネーションが信頼できないので、折り丁記号(signature)を使用することとなった。作者や特定の植字工に固有の綴りや単語、総数700についてテキストを検索し、折り丁記号(ページに相当)毎のカウントを行ない、その出現パターンを調べた。その結果、予想通りにパターンが確認され、研究としては成功をしたものと評価される。[3]

【評価・反省】

《プログラムに使用した言語について》

現在と異なり、研究開始時点では一般の英語関係の研究者が手軽に使用できる言語はBASICであった。同分野において電算機を使用している研究者が全て望んでいたのは、言葉の切れ目を認識して単語に切りわけてくれるなどの関数であったが、手に入る言語では、そのような関数がない為に、そのルーチンを書くだけでも試行錯誤を繰り返して貴重な時間を費やしてしまった。現在使用でき

るプログラム言語でも、英語英文学、文献学で必要とされる関数は当然のことながら標準となっていない。やはり、電「算」機のままであり、この研究会の進む道は遠いと思われる。

【入力したファイルの今後】

今回の『妖精の女王』に限らず、入力済みのテキストデータは、その殆どが個人の所有・管理下に置かれているものであり、研究者の死亡などで一代限りのデータとなる可能性を持つものが多い。このことは、英語に限らず他の言語を使用したテキストデータにもあてはまるものである。かなりの数の言語テキストが電算機可読形式媒体に入力されているものと考えられるのであり、テキストデータの統合化が強く望まれている。しかし、各言語毎のデータベース化は、余りにも容易な解決方法であり、データベースというよりは、単にデータバンクとして倉庫代わりの使用方法となってしまう、各国語での翻訳作品の占める位置が曖昧になるなどの弊害も予想される。今日の急務となっているのは、大規模なテキストデータ所在確認調査を行ない、他のインド・ヨーロッパ諸語のテキストデータとの整合性を図ることではなかろうか。テキストデータの整合性を欠いていれば、どのように大きなテキストデータベースも単に大きな本と同じ意味しか持たないものとなるだろう。

注

[1] 『言語研究の中の計算機』計算機利用言語学研究会編 2号 1984

[2] Sato, Haruo: On the Possibilities of Microcomputers in Literary Research -- A Memorandum for Processing The Faerie Queene, Proc. Nihon Univ. Sch. of Dent., 1982

[3] Yamashita, Hiroshi: The Printing of The First Part (Books I-III) of The Faerie Queene in 1590 --(I), Studies in Languages and Cultures, Univ. of Tsukuba, 13, 1982 参考資料-1

参考資料 - 1
 植字工同定用データ見本 筑波大学 山下 浩 論文より

KEY NO. 60 CLASSIFICATION NO. 5 KEY = EIES

20 a

EIES																	
		1	2	3	4	5	6	7	B	1	2	3	4	5	6	7	B
		R	V	R	V	R	V	R	V	R	V	R	V	R	V	R	V
1A																	
1B																	
1C	1																
1D		1	1	1	1	1											
1E																	
1F																	
1G																	
1H																	
1I																	
1K																	
1L																	
1M																	
1N																	
1O																	
1P																	
1Q																	
1R	1																
1S		1															
1T			1														
1V																	
1X																	
1Y																	
1Z																	
2A																	
2B																	
2C																	
2D																	
2E																	
2F																	
2G																	
2H																	
2I																	
2K																	
2L																	
2M																	
2N																	
2O																	
2P																	
2Q																	

KEY NO. 64 CLASSIFICATION NO. 5 KEY = EYES

20 b

EYES																	
		1	2	3	4	5	6	7	B	1	2	3	4	5	6	7	B
		R	V	R	V	R	V	R	V	R	V	R	V	R	V	R	V
1A																	
1B																	
1C																	
1D																	
1E																	
1F																	
1G																	
1H																	
1I																	
1K																	
1L																	
1M																	
1N																	
1O																	
1P																	
1Q																	
1R																	
1S																	
1T																	
1V																	
1X																	
1Y																	
1Z																	
2A																	
2B																	
2C																	
2D																	
2E																	
2F																	
2G																	
2H																	
2I																	
2K																	
2L																	
2M																	
2N																	
2O																	
2P																	
2Q																	