

## 電子計算機による英語教科書 の語彙表作成について<sup>1)</sup>

兵庫県山崎高等学校

河 田 孝 義

広島大学理学部

友 枝 謙 二

広島大学総合科学部

三 浦 省 五

### はじめに

英語教育学という学問を志向する動きが見られて約10年を経る。そこでは、英語教育によって学習者はいかに変容していくか、教授内容は、いかに教授法と関連し合って学習者に作用しているかなど、英語教育に係る問題点を客観的に観察し、その法則性を発見することが学問的研究の第1歩と考えられる。

英語教育における中心的な位置を占める教育内容(教材)は、音声、文型・文法事項、語彙と用法、文化内容などのレベルからとらえることが普通であり、たとえば、語彙の研究を例にとっても、より好ましい教材開発を目指して、各種の語彙選定や分析が行なわれ、多くの資料が提供されてきた。Thorndikeの1万語表、Basic English、Palmerの3千語表、Westの語の意味別頻度表、最近では、Carrollらの語彙分析、わが国でも、全英連の「高校基本英語活用集」、速川浩氏の「教科書に現われた英語単語の研究」や「中学校英語教科書に見る単語・イディオム・文型の統計研究」、その他、紀要の類にも語彙の統計的研究が見られ、分析する資料は主観的に選択するにしても、分析方法は、客観的方法がとられているようである。

わが国の教科書の語彙分析は、語の頻度の記述のみに終わっているものが多く、その用法、意味の視点からの研究が比較的少ない。したがって、語の頻度のみならず、用法、意味の参照も可能な基礎資料と呼ぶべきものが必要となり、昭和46年より電子計算機による語彙表の作成に着手した。ここでは、その手順のみを述べてみたい。

### 1. 語彙表の必要・十分条件

分析の対象となる英語教材に現われた語彙のentryの頻度だけではなく、意味と用法の観点からも語彙研究が行なえるように、その単語が使用されている英文も併記し、文意が不明瞭な場合は、教材そのものを参照できるように、その単語が現われている箇所をも明示している語彙表であることが望ましい。

### 2. 分析のための教科書

昭和50年度には、高等学校用英語B読本約20種類のうち、T社の*New Horizon English*

<sup>1)</sup> 本研究は、昭和50年度および51年度文部省科学研究費(特定研究・科学教育)補助金により、垣田直己教授を研究班長として行なわれたものの1部である。

Readers, S社の *Highroad to English Reading*, K社の *New Age Readers*, 51年度には, 中学校用教科書4種類, C社の *Everyday English Revised*, T社の *New Horizon English Course Revised Edition*, K社の *New Prince English Course*, S社の *Total English Junior Crown Series Revised Edition*, の語彙表を作成し, 中学校用については, 4つの語彙表をさらに1つにまとめて, 比較を容易にした。以下主として中学校用教科書に関して, 入力データの作成とプログラム開発について述べる。

### 3. プログラム開発と入力データの作成

語彙表作成のプログラムは, 1の必要・十分条件を満足させ, かつ, 入力データの作成および修正が容易であるように設計された。

入力データはパンチカードに穿孔された, 入力データの成が容易となるように, 次の符号を除いては, 全て原文(入力する教科書の英文)通りに穿孔した。

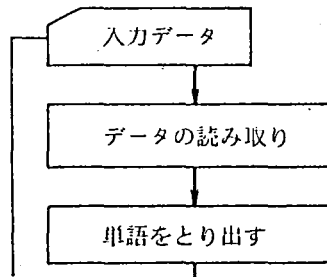
---

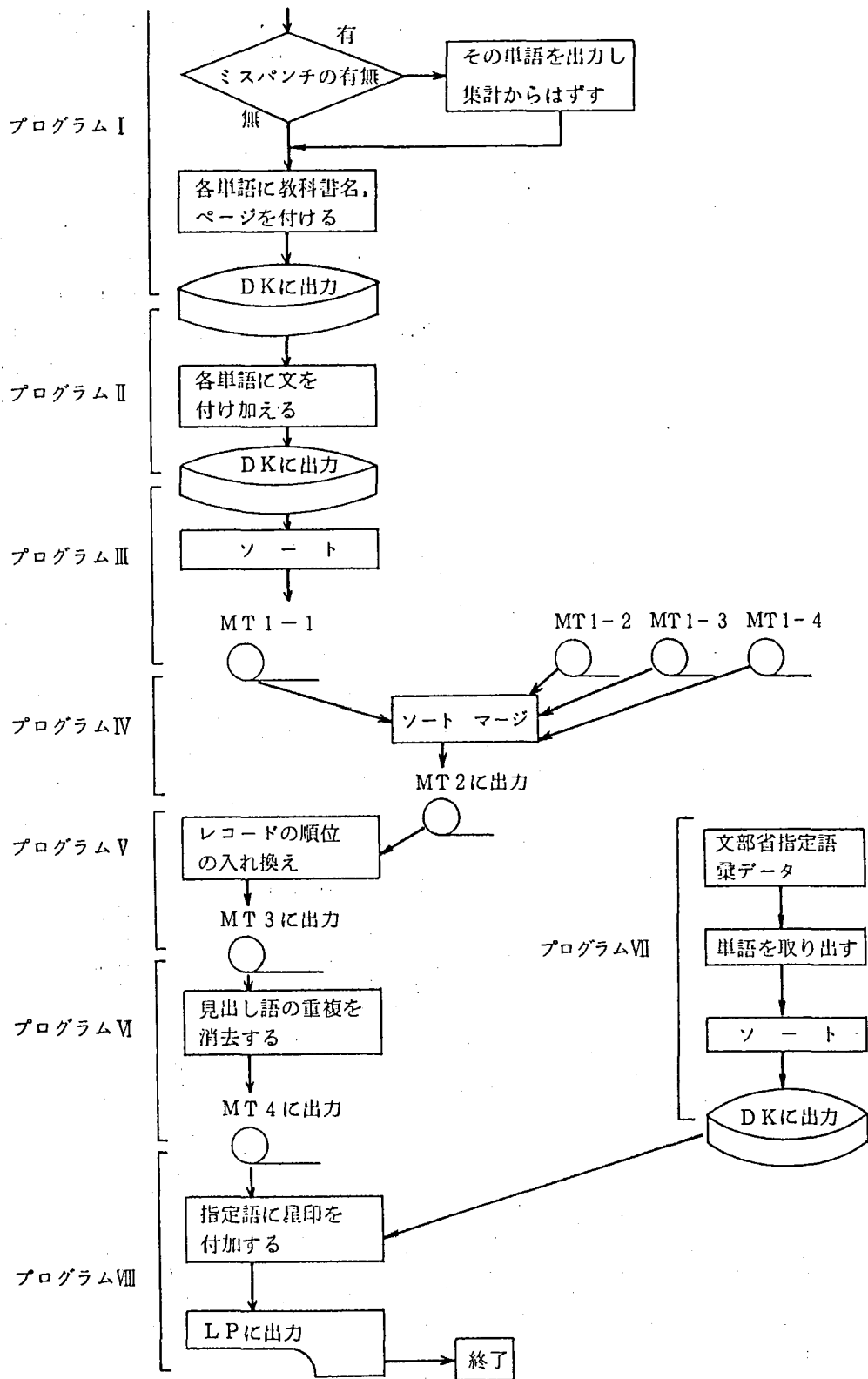
引用符( " ")	→	'
疑問符( ? )	→	*
アポストロフィ( ' )	→	'
S + アポストロフィ( S' )	→	S'
感嘆符( ! )	→	**
コロン( : )	→	..
セミコロン( ; )	→	..

---

ただし, 引用符, 疑問符, コロン, セミコロン, 感嘆符, ピリオド, コンマの前後は1カラム以上空白とした。これらの特殊記号を変更した理由は, 主としてHコードの穿孔機を用いたため, その制約に帰因している。また, 教科書の本文の前には, たとえば, ¥¥NH-3のように, *New Horizon English Course*の3学年用を表わすカードを加え, 各ページの最初には, たとえば, ¥P 25と25ページを表わす記号を穿孔した。

プログラムは, 次のフローチャートが示すように, 大きく8つの段階にわかれている。





各プログラム毎に、*Everyday English Revised*, 2学年用, 9ページのTom: Does your uncle play tennis? という文が、どのように処理されるか述べてみる。このデータは、次のような形で読み込まれる。(EEは *Everyday English* の略)

---

¥¥EE-2 ¥P9 TOM .. DOES YOUR UNCLE PLAY TENNIS \*

---

プログラムⅠでは、読み込まれた文は、単語ごとにばらばらにされ、見出し語となる。各々の単語は、穿孔のミスはないかチェックされる。つづり字のミスのチェックではなく、プログラムによって禁止されている文字の混入をチェックするのである。たとえば、ENGLISHのように数字が混入していたり、……PLAY TENNIS\*のように、疑問符の前に1カラム以上の空白が無い場合を対象としている。もし、誤りがあれば、集計からはずされ、正しい単語については、その後に教科書名およびページが付加される。この結果は次の通りである。

---

TOM	EE-2	9
:	EE-2	9
DOES	EE-2	9
YOUR	EE-2	9
UNCLE	EE-2	9
PLAY	EE-2	9
TENNIS	EE-2	9
?	EE-2	9

---

このようにして、コロンや疑問符をも含めて、すべてのシンボルは、プログラムⅡの入力データとなる。プログラムⅡの段階では、それぞれの見出し語と教科書名、ページ数の後に、見出し語の含まれている英文(ピリオドまたは疑問符までの文)を付加する。

---

TOM	EE-2	9	TOM : DOES YOUR UNCLE PLAY TENNIS ?
:	EE-2	9	TOM : DOES YOUR UNCLE PLAY TENNIS ?
DOES	EE-2	9	TOM : DOES YOUR UNCLE PLAY TENNIS ?
YOUR	EE-2	9	TOM : DOES YOUR UNCLE PLAY TENNIS ?
UNCLE	EE-2	9	TOM : DOES YOUR UNCLE PLAY TENNIS ?
PLAY	EE-2	9	TOM : DOES YOUR UNCLE PLAY TENNIS ?
TENNIS	EE-2	9	TOM : DOES YOUR UNCLE PLAY TENNIS ?
?	EE-2	9	TOM : DOES YOUR UNCLE PLAY TENNIS ?

---

このデータをディスクに出力し、ユーティリティプログラムを用いてソートする。ソートは、ピリオド、コンマなどの特殊記号に始まり、アルファベット、数字の順に並べられる。こうしてソートしたデータは次の通りであり、教科書の種類ごとに磁気テープ(MT 1-1~MT 1-4)に出力された。

?	EE-2	9	TOM : DOES YOUR UNCLE PLAY TENNIS ?
:	EE-2	9	TOM : DOES YOUR UNCLE PLAY TENNIS ?
DOES	EE-2	9	TOM : DOES YOUR UNCLE PLAY TENNIS ?
PLAY	EE-2	9	TOM : DOES YOUR UNCLE PLAY TENNIS ?
TENNIS	EE-2	9	TOM : DOES YOUR UNCLE PLAY TENNIS ?
TOM	EE-2	9	TOM : DOES YOUR UNCLE PLAY TENNIS ?
UNCLE	EE-2	9	TOM : DOES YOUR UNCLE PLAY TENNIS ?
YOUR	EE-2	9	TOM : DOES YOUR UNCLE PLAY TENNIS ?

こうして作成した4本のMTには、それぞれの教科書ごとに上記のフォーマットで、単語と英文が収められており、この4本を、語彙表の便宜を考慮してMT2にソートマージした。これがプログラムⅣである。“TENNIS”の項を例示すると下記の通りになっている。(NHは *New Horizon English Course* の略)

TENNIS	EE-1	56	DICK : HE PLAYS TENNIS EVERY DAY .
TENNIS	EE-1	92	I CAN PLAY TENNIS , TOO .
TENNIS	EE-2	9	TOM : DOES YOUR UNCLE PLAY TENNIS ?
TENNIS	EE-2	9	TOM : WHO PLAYS TENNIS BETTER , YOU OR YOUR UNCLE ?
TENNIS	EE-3	32	I PLAY EITHER TENNIS OR TABLE TENNIS AFTER SCHOOL .
TENNIS	EE-3	32	I PLAY EITHER TENNIS OR TABLE TENNIS AFTER SCHOOL .
TENNIS	NH-1	53	DO YOU LIKE TENNIS ?
TENNIS	NH-1	53	I LIKE TENNIS VERY MUCH .
TENNIS	NH-1	53	DOES YOUR BROTHER LIKE TENNIS ?
TENNIS	NH-1	53	HE LIKES TENNIS .
TENNIS	NH-1	54	JANE : HE LIKES TENNIS AND BASEBALL .
TENNIS	NH-1	54	HE PLAYS TENNIS EVERY SUNDAY .
TENNIS	NH-1	56	HE SOMETIMES PLAYS TENNIS WITH MY FATHER .
TENNIS	NH-1	61	JANE PLAYS TENNIS EVERY SUNDAY .
TENNIS	NH-1	61	SHE CAN PLAY TENNIS VERY WELL .
TENNIS	NH-1	62	WE PLAY TENNIS IN SPRING AND FALL .
TENNIS	NH-1	62	CAN YOU PLAY TENNIS ?
TENNIS	NH-1	62	I PLAY TENNIS EVERY SATURDAY .
TENNIS	NH-1	63	KEN : DO YOU LIKE TENNIS , MIKE ?
TENNIS	NH-1	63	MIKE : I PLAY TENNIS .
TENNIS	NH-1	71	HE SOMETIMES PLAYS TENNIS WITH ME .
TENNIS	NH-1	83	LET'S PLAY TENNIS .
TENNIS	NH-2	2	MY FATHER SOMETIMES PLAYS TENNIS WITH HIM .
TENNIS	NH-2	12	I USUALLY PLAY TENNIS WITH HELEN ON SATURDAY .

このようにして、中学校用教科書をすべてソート、マージしたものは、ユーティリティ・プログラムを使用したため、アルファベットの前に特殊記号が現われている。プログラムⅤでは、アルファベットを先頭に出し、数字、特殊記号の順になるように入れ換えた。プログラムⅥでは、見やすさという点から、重複した見出し語は、最初のものを残して、すべて消去した。また、プログラムⅦとⅧで文部省中学校学習指導要領に示されている必修語彙に星印を付しラインプリンターで出力した。その1部は次の通りである。

PLANTS	TE-2	8	' I WANT TO TALK TO YOU ABOUT PLANTS .
	TE-2	8	TODAY YOU CAN SEE MANY KINDS OF PLANTS .
	TE-2	10	YOU MAY LOOK AT THE PLANTS AND HAVE LUNCH AT THE GARDENS .
	TE-2	10	THEY SEE MANY STRANGE PLANTS .
	TE-2	10	THERE ARE JUNGLE PLANTS AND WATER PLANTS .
	TE-2	10	THERE ARE JUNGLE PLANTS AND WATER PLANTS .
	TE-2	12	YOU MAY LOOK AT THE PLANTS .
	TE-2	32	YOU MUST STUDY THE ANIMALS AND PLANTS AROUND YOU LIKE HIM ' .
	TE-3	18	' BUT YOU SHOULD STUDY PLANTS AND THE EARTH , TOO ,
	TE-3	20	' BECAUSE THERE ARE ONLY A FEW PLANTS AND ANIMALS IN ENGLAND ' , SAID THE PROFES SOR .
	TE-3	20	AND SO IF YOU WANT TO KNOW THE PLANTS AND ANIMALS OF THE WORLD , GO TO SEE THEM YOURSELF .
	TE-3	24	THE JUNGLE WAS FILLED WITH STRANGE PLANTS AND BEAUTIFUL FLOWERS .
PLAY	EE-1	69	I PLAY WITH MY FRIENDS .
	EE-1	67	KATE : PLAY THE PIANO , GRACE .
	EE-1	91	I CAN PLAY BASEBALL .
	EE-1	92	I CAN PLAY BASEBALL .
	EE-1	92	I CAN PLAY TENNIS , TOO .
	EE-1	92	CAN YOU PLAY BASEBALL , JIRO ?
	EE-1	92	I PLAY IT AFTER SCHOOL .
	EE-1	93	YOU MAY PLAY IT , JIRO .
	EE-2	9	TOM : DOES YOUR UNCLE PLAY TENNIS ?
	EE-2	28	PLAY THE GUITAR , JANE ' , SAID KATE .
	EE-2	33	ORPHEUS WANTED TO PLAY IT .
	EE-2	34	ORPHEUS BEGAN TO PLAY THE HARP .
	EE-2	62	JIRO : ARE YOU GOING TO PLAY IN THE GAME ?
	EE-2	64	EVEN NANCY KNOWS HOW TO PLAY IT .
	EE-3	3	SOON THE GREAT ORGAN STARTED TO PLAY .
	EE-3	32	I PLAY EITHER TENNIS OR TABLE TENNIS AFTER SCHOOL .
	EE-3	61	I HAD FEW FRIENDS , AND I WANTED MY MOTHER TO PLAY WITH ME .
	EE-3	62	AND I SAID TO HER , ' MOTHER , YOU MUST BE GETTING BETTER I CAN YOU COME OUT AND PLAY WITH ME ?
	EE-3	69	LONG AGO PEOPLE THOUGHT THAT WITCHES CAME OUT TO PLAY THAT NIGHT .
	EE-3	89	BUT YOU LIKE WORK BETTER THAN PLAY , DON'T YOU ?
	EE-3	96	THE PIPER WALKED OUT INTO THE STREET AND BEGAN TO PLAY HIS PIPE .
	EE-3	97	WHEN HE BEGAN TO PLAY HIS WONDERFUL PIPE , ALL THE CHILDREN IN THE TOWN CAME RUN NING TO HIM .
	NH-1	61	I CAN PLAY BASEBALL .
	NH-1	61	I CAN PLAY BASEBALL , TOO .
	NH-1	61	SHE CAN PLAY TENNIS VERY WELL .
	NH-1	61	I SOMETIMES PLAY WITH HER .
	NH-1	62	WE PLAY TENNIS IN SPRING AND FALL .
	NH-1	62	CAN YOU PLAY TENNIS ?
	NH-1	62	I PLAY TENNIS EVERY SATURDAY .
	NH-1	62	CAN TOM PLAY BASEBALL ?
	NH-1	62	HE CAN PLAY BASEBALL VERY WELL .
	NH-1	63	MIKE : I PLAY TENNIS .
	NH-1	63	I SOMETIMES PLAY BASEBALL .

すべてのプログラムは、フォートランで作成され、全プログラムの長さは、64 KW ( 1W=32 bits ) である。使用電子計算機はHITAC-8700で、4種類の中学校用教科書12冊を8つのプログラムを用いて処理するために、76分 ( CPU時間 ) を要した。プログラムⅠ、Ⅱ、Ⅲでは、1種類の教科書 ( 3冊 ) を処理するために約10分、プログラムⅣで1分、プログラムⅤ、Ⅵではそれぞれ10分、ⅦとⅧでは約15分を要している。

#### 4. 展 望

教材内容ができる限り迅速にかつ正確に把握できるような語彙表作成のプログラムを概説してきたが、問題点として次のようなものがあげられる。1) ラインプリンターがすべて大文字であるので、たとえば、人名や月のMARCHと動詞のMARCHの区別ができないこと。2) 電子計算機を使用して意味論的範疇による語彙表作成は、不可能ではないにしても前途多難であること。3) 特殊記号の入力に再考を要することがあげられるが、最大の長所として膨大な量の資料を短時間に処理することができ、また、人力による資料作成よりもはるかに信頼性の高い資料をそろえることができるという点が考えられる。