An application of KWIC system to the processing of newspaper vocabulary data : Programs were written in COBOL

新聞用語調査の用例印字プログラム "COBOL—KWIC"

石 綿 敏 雄

0. はじめに

この稿は現在電子計算機を用いて行なっている新聞用語調査システムのなかで使用されている，用例印字プログラム，"COBOL—KWIC" について報告するものである。このなかには新聞用語調査の主要な流れの一つである，長単位処理のプログラムの一部も含まれているので，この稿はその長単位プログラムについての報告を兼ねてもいる。この長単位処理プログラムは，さきの1紙1年分の報告（報告37）のあとをうけて2紙1年分の処理に使用された。またこの COBOL—KWIC は新聞用語調査と同じフォーマットをもつものに対しては，それ以外のデータについても用例つきの総索引を作成する能力をもち，そように使用することが可能である。

1. プログラムのあらまし

はじめに，このプログラム作成の目的について述べたい。

現在行なっている新聞用語調査のプログラムでは，同語異語の操作が行なわれていない（国語研究報告37，3ページ）。また漢字の読みについても，原文について読んだのではなく，できあがった語い表について読んだので，原資料のなかでの実態を示していない面がある。たとえば表記形が同じ「いき」であれば，「行きと帰り」の「いき」でも，「粋」の「いき」でも，「行きます」の「いき」でも，「活きます」の「いき」でも，すべて同一語形としてまとめて

カウントしてしまっている。語い調査本来の目的からすればこれは分けてカウントしなければならないはずである。また意味用法が同じでも表記形式が異なれば異なった語として度数をカウントしている。たとえば「いきます」「行きます」の「いき」に当たる部分で，このばあい意味用法が全く同じであっても，表記された形式が異なるから，別の語としてカウントされるのである。「間」も「かん」か「あいだ」かは，複合語の一部でなく単独で現われたものについては，全体については，作業上どちらかにまとめておくという程度の処置をするほかなかった。これらの処置を適切にするためには原文のなかでの意味用法，原文のなかでの漢字のよみ方がわかるようにならなければならない。

新聞用語調査の長単位処理，短単位処理のプログラムの操作をへたあとで得られる語い表は以上のような性格をもつものである。この語い表は語い表というよりは同形文字列表というべきである。これは用語調査とすると大きな欠陥であるといわなければならない。この欠陥をカバーすべくなんらかの手を打たなくてはならない。そのためには，これらの語形をもとの文脈のなかに置いて考えればよいのであるが，その方法にはいろいろある。まずはじめから，漢字のよみ，同語異語のおそれのあるものはそれをつけて入力することである。これは徹底して行なえば完全な調査ができるはずであるが，いろいろ問題がある。プリエディトが大変だし，それをさん孔するのも量が多いのと処理やフォーマットを誤まりやすいので大変である。入力前および入力時に非常な負担がかかるのは一考を要しよう。次にいわゆる総索引式に出典表をつくり，それによっていちいち原文を捜して考えるという方法もある。これは入力は原文のままですむから容易であるが，いちいちもとのものを捜す手間が大変である。そこで最後に考えられるのが KWIC (Key Word in Context) をつくって，用例集を印字してしまうのである。これは入力も容易，捜す手間も省けるので，非常に能率的である。用語調査を電子計算機で行なうばあい，これはまず第一にとにかく有効な手段であるということができる。これなら漢字の読みあやまりもそれを利用したあとでのチェックでなくすことができるし，同語異語の操作も，人間の目と手によってではあるが，完全に行ないうるからである。しかも

後に述べるように，KWICのデータ排列に工夫をこらしておくことによって，作業が一括して行なうことができるようになり，機械と人間をまじえたトータル・システムとして，非常に有効なものであるということができよう。このように，KWICは用語調査システムとして有効なものであり，われわれの新聞用語調査についてみても，その欠陥をカバーすることができる方法としてはほとんど唯一のものであるということができるのではないか。

さて，KWICが用語のカウント調査について有効な方法であることを述べたが，その利点はそれだけ，つまり同語異語の判別だけにとどまるのではない。これは，ある語を中心として，それをとりまく前後の単語を整理した用例集として使用することができるように，プログラムを作成することができるから，そのようにすれば単語についての連続状況，たとえばその確率についても計算でき，語の意味用法についての記述に非常に有益な資料を提供し，文法的な用語についてのそれもできるので，文法の研究にも有益な資料を提供することができる。語いや文法の研究という，純粋に言語学的な問題へ資料を提供するのであるから，同時われわれがめざしている computational linguistics にとっても有益な資料を提供するということができよう。KWICは言語情報処理の研究開発のためにも有益なのである。さいわい，新聞用語調査の規模は300万語に達するので，この用例印字プログラムを作成すればこの300万語の用例集をつくることができる。これは日本語研究にとって非常に重要な資料であるということができよう。

このような意味のほかにKWICのプログラムをつくる意図はもう一つあった。それは，新聞用語調査の全データの1/3の量にあたる，いわゆる1紙1年分の処理にあたって使用した長単位処理プログラムの作成者が退職し，その取り扱いに困難な点を生じたので，その部分をつくり直すという意味があった。せっかく作り直すならば，ついでに新聞用語調査の入力フォーマットから多少はずれたものでも，これを受け入れて，総索引作成のために使用することができれば便利である。ただしあまりフォーマットを自由にすると，チェック機構として弱くなるので，それもほどほどにしなければならないが。またせっかく作り直すのだからKWIC自体にも便利なようなファイルをどこかで作っておき

たい。このようなことができれば，長単位の一部のシステムの作り直しをしても，十分その意味がある。そういうことで，入力フォーマットのチェックを多少ゆるめ，従来プログラム上直接は作られなかったファイル（同じ文中での同じ単語の数を足してしまうのでなく，文中の単語番号を記入したファイル）を作成することを，そのプロセスのなかに組みこんだ，システム・アナリシスをして，従来作成した語い調査プログラムの完全な実行を期しもしたのであった。

用例印字をKWICの形で行なうのは，大量のデータの印字として有効であり，用語の意味，用法の研究の上からも便利であり，それは同語異語の判別のときにも有効である。印字形式をKWICの形で行なうことにしたのはそのためである。またこのプログラムは，ごく一部のものを除いて他はすべてCOBOLを使用したので，このプログラム全体の名称を COBOL—KWIC とした。プログラム言語としてCOBOLを用いたのは，計算機使用の全体がすでにコンパイラの時代にはいっていることを考え，また他機種との互換性を考えたからである。それから， HITAC 3010 は，記憶装置の構造や命令語の体系が，全体として記号列の処理，たとえば自然語の処理にかなりよく適合している部分があるために，どうしてもアセンブラにたよることになりやすい。これでは機種の間の互換性のあるプログラムをつくることはできない。そこで，コンパイラでこのような問題をどのように解くことができるかを考えてみることも必要である。国語研究所の現在保有するプログラムのほとんどはアセンブラで書かれている。この COBOL—KWIC は，このような意味での完全 variable length dataの処理の，国語研究所のなかでのパイオニアの役目を果たすつもりもあった（例を示したのはその意味である）。それは一種の教育的な意味でのプロパガンダでもある。そこで COBOL—KWIC というような名前をこのプログラムに，あえて，つけてみたのである。COBOL を用いることに，さらに別な意味があることは本稿の７．問題点の項でも述べる。

では次に，プログラム・システムの構造の概略について述べたい。

以上に述べたような目的を果たすために，現在国語研究所に設置されている機種をどのように活用したらよいだろうか。また，現在あるデータをどのよう

に活用したらよいだろうか。

現在できている入力は紙テープからはいった原文であり，これは新聞にあったとおりの表記で，漢字かなまじり文である。漢字にかなはついていない。ところで，これをそのまま漢テレで印字することを考えてもよいのであるが，漢テレの印字速度が遅く，とても大量の用例を一時に打出すことはできない。そこで，ラインプリンタを利用することが考えられる。ラインプリンタを利用すると，漢字の情報をそのままとり出すのはやや困難であるが（方法がないことはない。後述），その程度のものでもかなり役立つだろう。このように考えると，漢テレ入力，ラインプリンタ出力という線がつよくなってくる。そこでこのようなプログラム構成を，COBOL—KWIC では考えることにしたのである。

この，漢テレ入力，ラインプリンタ出力という線は，高速漢字プリンタがないために行なうわけで，この状態においてはきわめて有益であるが，もし漢字の高速プリンタが存在するならば，その価値が大いに減ずることはいうまでもない。入力した形の，漢字かなまじり文で，そのまま出力できれば，その方がずっと有益であることは，いうまでもない。漢字プリンタを使用するときは，江川清氏が作成したプログラムなどを使用することもできよう。

しかしもし漢字高速プリンタが存在しても，ラインプリンタ出力が全く価値を失うかというと，そうはいえないのであって，たとえばかな漢字変換などのプログラムには，このようにしてよられたKWICが一番重要な研究資料を提供することになるだろう。語の用法や意味の研究，文法的な研究ならばこれをつかってじゅうぶんにできるはずである。だから，ラインプリンタ出力のKWICはいつまでも，その存在価値を保有するに違いない。

ところで，漢テレ入力，ラインプリンタ出力となると，ラインプリンタはかな専用であり，入力原文は原文どおりの漢字かなまじり文字なので，どうしても漢字解読を行なわなければならなくなる。しかも，研究のために使うのだから，なるべく解読正解率を高めたい。また，現在の新聞用語調査の判定に直接役立つことも重要な目的の一つにしている。そう考えて，この両者を解決できるように，新しい漢字解読プログラムを設計した。漢字の解読の正解率がよ

く，しかも同語異語の解明にとってつごうがよいという条件を一挙に解決できたのはまことにさいわいであったが，これは短単位処理のなかで，非常に便利なファイルができていたので，それを利用することによって解決したのである。

このように，COBOL—KWIC は，紙テープ入力データの各種のチェック，漢字解読，およびＫＷＩＣ本来の用例印字のプログラムからなっている。そのほかこれに付随して，辞書のメインテナンスのルーチン，用例の検索ルーチンがある。辞書のメンテナンスは，COBOL—KWIC システムが独自に使う辞書を作成しあるいはメンテナンスを行なうシステムであり，用例検索は，欲する用語の用例集を任意に作成することを可能にするプログラムである。

この COBOL—KWIC を構成するプログラムは全体で20数本ある。それは次の５つのグループに分けられる。

Ａ．チェック・プログラム・グループ。

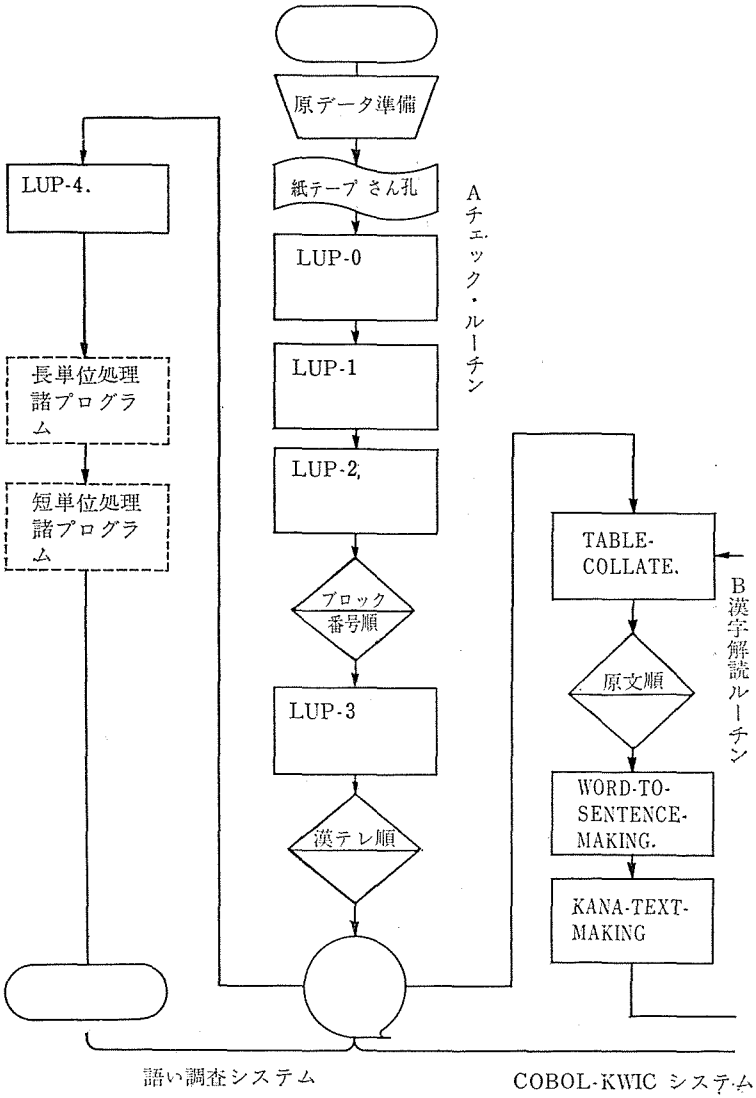
Ｂ．漢字解読プログラム・グループ。

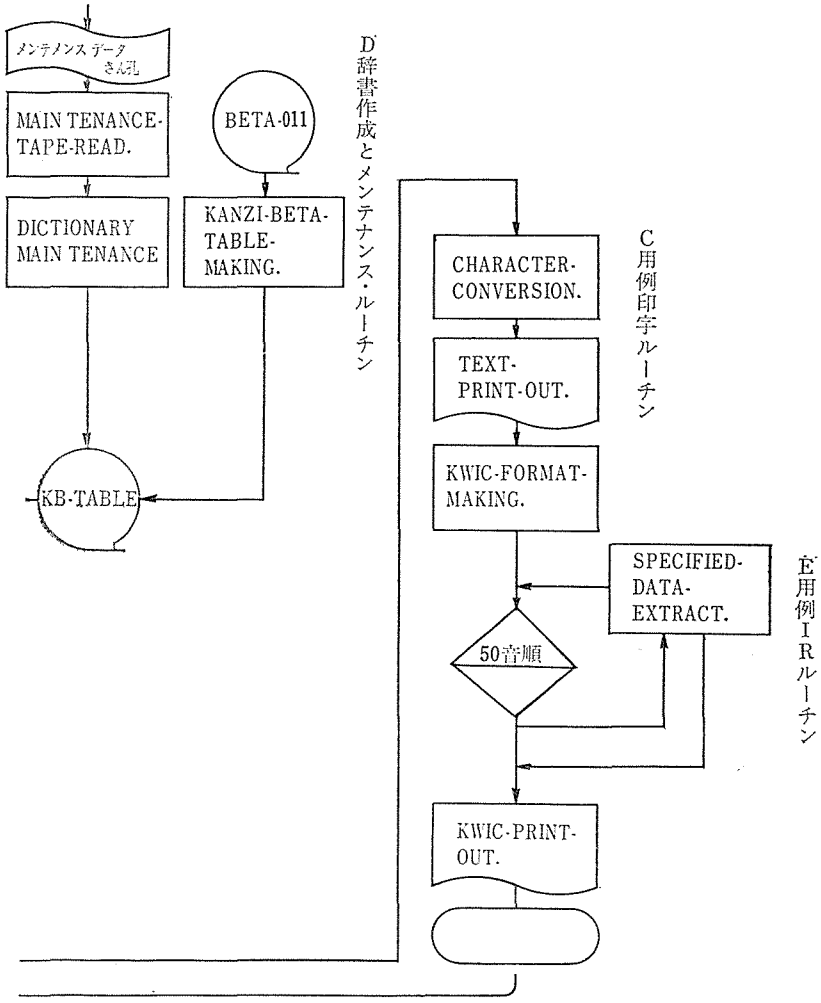
Ｃ．用例印字プログラム・グループ。

Ｄ．辞書の作成とメンテナスのプログラム・グループ。

Ｅ．用例検索（IR）プログラム。

このプログラムのシステム・アナリシスに当たって，全体をややこまかく分けることにした。これはプログラムを細分化することによって，いわば一種のモジュール化を行ない，その間に，必要に応じて自由に他のプログラムを挿入することができ，あるいは削除短絡することができるようにしたためである。このため，Ｂ以降のデータについては特別のものを除いてラベルを共通にし，フォーマットもできるだけ同じようにした。KANA-TEXT-MAKING のごとき，実はそう長いプログラムでないので他のプログラムにあわせてもよく，かつ辞書自体のなかに漢字によみがなをつけた形でなくて，かなだけのものをつけておいてもよいはずである。そうすればKANA-TEXT-MAKING は不要になる。にもかかわらずこれをつけておいたのは次のような考え方をしたからである。このKANA-TEXT-MAKING と入れかえにかなのうしろに漢字の情報を入れておくようにすれば，漢字情報をそえた印字がこのプログラムを延長す


原データ準備
紙テープ さん孔
LUP-0
LUP-1
LUP-2,
ブロック番号順
LUP-3
漢テレ順
Aチェック・ルーチン
LUP-4.
長単位処理諸プログラム
短単位処理諸プログラム
TABLE-COLLATE.
原文順
WORD-TO-SENTENCE-MAKING.
KANA-TEXT-MAKING
B漢字解読ルーチン
語い調査システム
COBOL-KWIC システム


メンテメンスデータ
さん孔
MAIN TENANCE-
TAPE-READ.
BETA-011
DICTIONARY
MAIN TENANCE
KANZI-BETA-
TABLE-
MAKING.
KB-TABLE
D 辞書作成とメンテナンス・ルーチン
CHARACTER-
CONVERSION.
TEXT-
PRINT-OUT.
KWIC-FORMAT-
MAKING.
C 用例印字ルーチン
SPECIFIED-
DATA-
EXTRACT.
50音順
KWIC-PRINT-
OUT.
E 用例IRルーチン

ることによっても可能になるからである。たとえばコウコウというのにはいろいろな漢字が当たるが，国語研漢テレコードによって印字すれば，コウ(A7)コウ(A1)によって「高校」を表わし，コウ (13) コウ (&E) によって「孝行」を表わし，コウ（1.）コウ (&E) によって「航行」を表わすこともできる。このようになっていれば大変便利だろう。フォーマットはコウコウ（A7 A1）でもかまわない。このようにすることはKANA-TEXT-MAKING のなかをほんの少し手入れするだけで，できるのである。これは一つの例であって，このような自由なプログラム上のさしかえができるようにするために，プ

```
06/01/70206543211   F NNS    LUP-2.                                  6
000100 IDENTIFICATION DIVISION.
000110 PROGRAM-ID. LUP-2.
000200 REMARKS. THIS PROGRAM MAKES THE ONE WORD RECORD
000210     OF LONG UNIT, FROM THE NEWSPAPER SAMPLING BLOCK.
000300 ENVIRONMENT DIVISION.
000400 CONFIGURATION SECTION.
000500 SOURCE-COMPUTER. 3010, MEMORY SIZE 20000 CHARACTERS,
000600     SIMULTANEOUS-CONTROL, PAPER-READER, PRINTER 1.
000700 OBJECT-COMPUTER. 304, MEMORY ADDRESS 1000 THRU 19999,
000800     NO ROLLBACK.
000900 SPECIAL-NAMES. INTERRUPT-BUTTON, ON STATUS IS BUTTON-ON.
001000 INPUT-OUTPUT SECTION.
001100 FILE-CONTROL.
001200     SELECT INPUT-FILE, ASSIGN TO TAPES 1 FOR MULTIPLE REEL.
001300     SELECT OUTPUT-FILE ASSIGN TO TAPES 3 FOR MULTIPLE REEL.
001400     SELECT SPECIAL-FILE ASSIGN TO TAPES 4 FOR MULTIPLE REEL.
001500     SELECT PRINT-FILE, ASSIGN TO PRINTER 1.
001600 I-O-CONTROL.
001700     RERUN ON 2, EVERY END OF REEL OF SPECIAL-FILE.
001800 DATA DIVISION.
001900 FILE SECTION.
002000 FD  SPECIAL-FILE, RECORD CONTAINS 10 CHARACTERS,
002100     LABEL RECORD IS OMITTED,
002200     DATA RECORD IS SPECIAL-AREA.
002300 01  SPECIAL-AREA.
002400     02 USELESS, PICTURE IS X(10).
002500 FD  PRINT-FILE, RECORD CONTAINS 120 CHARACTERS,
002600     LABEL RECORD IS OMITTED,
002700     DATA RECORD IS PRINT-AREA.
002800 01  PRINT-AREA.
002900      02 BLOCK-NUMBER, PICTURE IS X(7).
003000     02 FILLER,       PICTURE IS X(2).
003100      02 SENTENCE-NUMBER, PICTURE IS X(2).
003200      02 FILLER,       PICTURE IS X(2).
003300      02 WORD-NUMBER,  PICTURE IS X(3).
003400      02 FILLER        PICTURE IS X(2).
003500     02 WORD.
003600      03 LETTER, PICTURE IS X(2) OCCURS 20 TIMES.
003620     02 MESSAGE, REDEFINES WORD.
003640      03 FILLER, PICTURE IS X(40).
003700      02 FILLER,       PICTURE IS X(62).
003800 FD  INPUT-FILE, RECORD CONTAINS 4018 CHARACTERS,
003900     LABEL RECORDS STANDARD,
004000     VALUE OF ID IS "LUD-1    ",
004100     ACTIVE-TIME IS 1,
004200     DATA RECORD IS INPUT-AREA.
```

ログラムの細分化を行なったのである。オペレーションのばあいめんどうではあるが，１ジョブが適当な時間に終わるので，その点にメリットがある。

ついでながら，漢字が（Ａ７Ａ１）のように表記されたばあい， decoding が問題ではある。これを人間が行なわなければならないからである。しかし“Ａ７”→「高」のような操作を簡単にできるような索引を同じくCOBOLで

```
004300 01  INPUT-AREA.
004400     02 BLOCK-NUMBER, PICTURE IS X(18).
004500     02 AREA-UNIT, OCCURS 1 TO 100 TIMES.
004600      03 LETTER, PICTURE IS X(2), OCCURS 40 TIMES.
004700 FD  OUTPUT-FILE, BLOCK CONTAINS 10 RECORDS,
004800     RECORD CONTAINS 60 CHARACTERS,
004900     LABEL RECORD ARE STANDARD,
005000     VALUE OF ID IS "LUD-2   ",
005100     ACTIVE-TIME IS 1,
005200     DATA RECORD IS OUTPUT-AREA.
005300 01  OUTPUT-AREA.
005400     02 SOURCE.
005500      03 BLOCK-NUMBER,     PICTURE IS X(7).
005600      03 SENTENCE-NUMBER, PICTURE IS X(2).
005700      03 WORD-NUMBER,      PICTURE IS X(3).
005800     02 STRATA,            PICTURE IS X(8).
005900     02 HEAD-WORD.
006000      03 LETTER, PICTURE IS X(2), OCCURS 20 TIMES.
006100 WORKING-STORAGE SECTION.
006200 01  COUNTERS.
006300     02 I-POSITION, PICTURE IS 9(3).
006400     02 O-POSITION, PICTURE IS 9(3).
006500     02 WORD-COUNTER, PICTURE IS 9(3).
006600     02 WORD-COUNT-DEPOT.
006700      03 C-1 PICTURE IS 9(1).
006800      03 C-2 PICTURE IS 9(2).
006900     02 WORD-COUNT-STORE.
007000      03 C-2, PICTURE IS 9(2).
007100      03 C-1, PICTURE IS 9(1).
007200     02 SENTENCE-COUNTER, PICTURE IS 9(2).
007300     02 I-AREA, PICTURE IS 9(3).
007400 01  WORK-AREA.
007500     02 I-BLOCK-NUMBER.
007600      03 FILLER, PICTURE IS X(3).
007700      03 I-1,    PICTURE IS X(1).
007800      03 FILLER, PICTURE IS X(1).
007900      03 I-2,    PICTURE IS X(1).
008000      03 FILLER, PICTURE IS X(1).
008100      03 I-3,    PICTURE IS X(1).
008200      03 FILLER, PICTURE IS X(1).
008300      03 I-4,    PICTURE IS X(1).
008400      03 FILLER, PICTURE IS X(1).
008500      03 I-5,    PICTURE IS X(1).
008600      03 FILLER, PICTURE IS X(1).
008700      03 I-6,    PICTURE IS X(1).
```

つくっておいた。これはかなり便利に使える。

COBOL でプログラムを作成したばあいに問題になるのは，後にも述べるように一般にアセンブラで書いたばあいより，処理時間が長くなることであろうと思う。いわば常識的なことであるが，大量のデータを，長い処理時間をかけて行なうとなんらかの手段をとらざるを得ない。そこで一番ふつうの方法はリランを行なうことだと思うが， COBOL-KWIC には独特の強力なリラン・ルーチンが組みこまれて威力を発揮している。それはオペレーターが欲するときに作業をやめて，次回再開したばあいに，まさにその次のデータからしごとを始めるという，「任意時リラン」である。このリランは，COBOL-KWIC のほ

```
008800       03 FILLER, PICTURE IS X(1).
008900       03 I-7,    PICTURE IS X(1).
009000       03 FILLER, PICTURE IS X(2).
009100     02 O-BLOCK-NUMBER.
009200       03 O-1, PICTURE IS X(1).
009300       03 O-2, PICTURE IS X(1).
009400       03 O-3, PICTURE IS X(1).
009500       03 O-4, PICTURE IS X(1).
009600       03 O-5, PICTURE IS X(1).
009700       03 O-6, PICTURE IS X(1).
009800       03 O-7, PICTURE IS X(1).
009900     02 LETTER.
010000       03 X-SIDE, PICTURE IS X(1).
010100       03 Y-SIDE, PICTURE IS X(1).
010200     02 TEMPORAL.
010300       03 LEFT-SIDE, PICTURE IS 9(1).
010400       03 RIGHT-SIDE, PICTURE IS 9(1).
010500     02 STRATA-STORAGE.
010600       03 G-AREA, PICTURE IS X(2).
010700       03 P-AREA, PICTURE IS X(2).
010800       03 S-AREA, PICTURE IS X(2).
010900       03 T-AREA, PICTURE IS X(2).
011000     02 SIGNAL, PICTURE IS X(1).
011100     02 STRATA-SIGNAL, PICTURE IS X(2).
011200 CONSTANT SECTION.
011300     77 EI-MARK, PICTURE IS X(1), VALUE IS "".
011400     77 ONE, PICTURE IS 9(3), VALUE IS 001.
011500     77 ONE-2, PICTURE IS 9(2), VALUE IS 01.
011600     77 NUMBER-SIGN, PICTURE IS X(2), VALUE IS "=#".
011700     77 BEGIN-PAREN, PICTURE IS X(2), VALUE IS "(".
011800     77 END-PAREN, PICTURE IS X(2), VALUE IS ")".
011900     77 WORD-LENGTH-MAX, PICTURE IS 9(3), VALUE IS 020.
012000     77 GO-TO-PRINT-ROUTINE, PICTURE IS X(1), VALUE IS "P".
012100     77 BEGIN-FLAG, PICTURE IS X(1), VALUE IS "B".
012200     77 G-STRATA, PICTURE IS X(2), VALUE IS "G".
012300     77 P-STRATA, PICTURE IS X(2), VALUE IS "P".
012400     77 S-STRATA, PICTURE IS X(2), VALUE IS "S".
012500     77 T-STRATA, PICTURE IS X(2), VALUE IS "T".
012600     77 DATA-ZERO PICTURE IS X(2), VALUE IS "Z".
012700     77 I-AREA-FULL, PICTURE IS 9(3), VALUE IS 040.
012800 PROCEDURE DIVISION.
012900 BEGIN-ROUTINE. OPEN INPUT INPUT-FILE, OUTPUT OUTPUT-FILE,
013000     PRINT-FILE AND SPECIAL-FILE.
013100 READ-INPUT-FILE.
013200     IF BUTTON-ON, GO TO RERUN-ROUTINE.
```

とんどすべてのプログラムに組みこまれている。

KWIC というと，日本語なら助詞，助動詞や副詞，接続詞など，英語なら冠詞や前置詞などは省かれて，見出し語とならないのがふつうである。しかし用語調査，言語研究のばあいは，それとは条件が異なり，そういうものこそ，まさに欲しいものである。そこで，この KWIC プログラムでは，そのような意味での単語の削除は一切行なっていない。この点一般につくられている KWIC とは，少し異なっているところがある。

上にのべたC群のなかのプログラムのソートで，直前語優先と直後語優先ということをのべたが，実際に得られた用例表でどのような相違があるかについ

```
013300      READ INPUT-FILE, AT END GO TO END-ROUTINE.
013400      MOVE ALL ZEROES TO SENTENCE-COUNTER.
013500      MOVE ALL ZEROES TO WORD-COUNTER.
013600      MOVE ONE TO I-AREA.
013700      MOVE ONE TO I-POSITION.
013800      MOVE ONE TO O-POSITION.
014000      MOVE ALL ZEROES TO STRATA-STORAGE.
014100 BLOCK-NUMBER-ADJUSTMENT.
014200      MOVE BLOCK-NUMBER IN INPUT-AREA
014300      TO I-BLOCK-NUMBER IN WORK-AREA.
014400      MOVE I-1 TO O-1.
014500      MOVE I-2 TO O-2.
014600      MOVE I-3 TO O-3.
014700      MOVE I-4 TO O-4.
014800      MOVE I-5 TO O-5.
014900      MOVE I-6 TO O-6.
015000      MOVE I-7 TO O-7.
015100 SYMBOL-EXAMINE.
015200      MOVE LETTER IN INPUT-AREA (I-AREA, I-POSITION)
015300      TO LETTER IN WORK-AREA.
015400      IF X-SIDE OF LETTER IN WORK-AREA IS EQUAL TO
015500      EI-MARK, GO TO READ-INPUT-FILE.
015900      IF LETTER IN WORK-AREA IS EQUAL TO BEGIN-PAREN,
016000      GO TO STRATA-PROCESSING.
016100      IF LETTER IN WORK-AREA IS EQUAL TO SPACE,
016200      GO TO WRITE-OUTPUT-FILE.
016300      IF LETTER IN WORK-AREA IS EQUAL TO NUMBER-SIGN,
016400      GO TO SENTENCE-COUNTING.
016410      IF X-SIDE OF LETTER IN WORK-AREA IS EQUAL TO EI-MARK,
016420      GO TO READ-INPUT-FILE.
016500      IF Y-SIDE OF LETTER IN WORK-AREA IS EQUAL TO EI-MARK,
016600      GO TO READ-INPUT-FILE.
016700      GO TO LETTER-MOVE.
017200 SENTENCE-COUNTING.
017300      ADD ONE-2 TO SENTENCE-COUNTER.
017400      MOVE ALL ZEROES TO WORD-COUNTER.
017500      PERFORM ADD-ONE-TO-I-POSITION THRU ADD-EXIT.
017510      IF LETTER IN INPUT-AREA (I-AREA, I-POSITION)
017520      IS EQUAL TO SPACE,
017600      GO TO NEXT-CHARACTER-POSITIONNING.
017610      GO TO SYMBOL-EXAMINE.
017700 LETTER-MOVE.
017800      IF O-POSITION IS GREATER THAN WORD-LENGTH-MAX,
017900      MOVE GO-TO-PRINT-ROUTINE TO SIGNAL,
018000      GO TO WRITE-OUTPUT-FILE.
```

て，少しふれておく。前優先で得られる用例表は，見出し語の直前の語がどんな単語であるかによってまず排列され，それが同じ語である範囲内において後の語が字母順に排列されるのである。たとえば動詞イルを見出し語として検索すると，「……ガ　イル」「……テ　イル」「……ニ　イル」などの動詞と助詞の結びつきがまずはっきりとつかめる。後ろ優先にしたばあいには，前優先のばあいと逆で，うしろの語をまず字母順に並べ，それが同じ語である範囲内において前の語を分類してあるのである。たとえば助詞ヲの見出しでみると，まずそのうしろの動詞が「……ヲ　アタエル」「……ヲ　イトナム」「……ヲ

```
018100      MOVE LETTER IN WORK-AREA TO LETTER IN
018200      OUTPUT-AREA (O-POSITION).
018300      ADD ONE TO O-POSITION.
018400      GO TO NEXT-CHARACTER-POSITIONNING.
018500 PRINT-ROUTINE.
018600      MOVE HEAD-WORD TO WORD IN PRINT-AREA.
018700      MOVE BLOCK-NUMBER IN OUTPUT-AREA
018800      TO BLOCK-NUMBER IN PRINT-AREA.
018900      MOVE SENTENCE-NUMBER IN OUTPUT-AREA
019000      TO SENTENCE-NUMBER IN PRINT-AREA.
019100      MOVE WORD-NUMBER IN OUTPUT-AREA
019200      TO WORD-NUMBER IN PRINT-AREA.
019300      WRITE PRINT-AREA.
019400      GO TO WRITE-REAL.
019500 NEXT-CHARACTER-POSITIONNING.
019600      PERFORM ADD-ONE-TO-I-POSITION THRU ADD-EXIT.
019700      GO TO SYMBOL-EXAMINE.
019800 WRITE-OUTPUT-FILE.
019900      ADD ONE TO WORD-COUNTER.
020000      MOVE WORD-COUNTER TO WORD-COUNT-DEPOT.
020100      MOVE C-1 IN WORD-COUNT-DEPOT TO C-1 IN WORD-COUNT-STORE.
020200      MOVE C-2 IN WORD-COUNT-DEPOT TO C-2 IN WORD-COUNT-STORE.
020300      MOVE O-BLOCK-NUMBER IN WORK-AREA
020400      TO BLOCK-NUMBER IN OUTPUT-AREA.
020500      MOVE STRATA-STORAGE TO STRATA IN OUTPUT-AREA.
020600      MOVE SENTENCE-COUNTER TO SENTENCE-NUMBER IN OUTPUT-AREA.
020700      MOVE WORD-COUNT-STORE TO WORD-NUMBER IN OUTPUT-AREA.
020800      IF SIGNAL IS EQUAL TO GO-TO-PRINT-ROUTINE,
020900      MOVE SPACE TO SIGNAL,
021000      GO TO PRINT-ROUTINE.
021100 WRITE-REAL. WRITE OUTPUT-AREA.
021200      MOVE SPACES TO HEAD-WORD.
021210      MOVE ONE TO O-POSITION.
021300      GO TO NEXT-CHARACTER-POSITIONNING.
021400 STRATA-PROCESSING.
021500      PERFORM ADD-ONE-TO-I-POSITION THRU ADD-EXIT.
021600      MOVE LETTER IN INPUT-AREA (I-AREA, I-POSITION)
021700      TO LETTER IN WORK-AREA.
021800      IF LETTER IN WORK-AREA IS EQUAL TO END-PAREN,
021900      GO TO NEXT-CHARACTER-POSITIONNING.
022000      IF Y-SIDE OF LETTER IN WORK-AREA IS NUMERIC,
022100      GO TO NUMERIC-RIGHT.
022200      MOVE ALL ZEROES TO TEMPORAL.
022600      GO TO SIGNAL-SET.
022700 NUMERIC-RIGHT.
022800      MOVE Y-SIDE OF LETTER IN WORK-AREA
022900      TO RIGHT-SIDE OF TEMPORAL.
```

イウ」「……ヲ　ウケル」「……ヲ　ウンエイスル」「……ヲ　エル」などと並び，しかも同じ動詞の前があとから字母順にならべられて，たとえばモツについては「イミ　ヲ　モツ」「カクシン　ヲ　モツ」「カチ　ヲ　モツ」「カンシン　ヲ　モツ」「キタイ　ヲ　モツ」「キノウ　ヲ　モツ」「キョウミ　ヲ　モツ」「コウカ　ヲ　モツ」「コウノウ　ヲ　モツ」「サヨウ　ヲ　モツ」「シュダン　ヲ　モツ」などのような連語構成が一覧できるようになる。これによってわかるように，前優先も後優先も，どちらも特徴があるの

```
023000      PERFORM ADD-ONE-TO-I-POSITION.THRU ADD-EXIT.
023100      MOVE LETTER IN INPUT-AREA (I-AREA, I-POSITION)
023200      TO LETTER IN WORK-AREA.
023300      IF Y-SIDE OF LETTER IN WORK-AREA IS NUMERIC,
023400      MOVE RIGHT-SIDE TO LEFT-SIDE,
023500      MOVE Y-SIDE OF LETTER IN WORK-AREA TO RIGHT-SIDE,
023600      GO TO NUMERIC-MOVE.
023700      PERFORM SUBTRACT-ONE-FROM-I-POSITION THRU SUB-EXIT.
023800      GO TO NUMERIC-MOVE.
023900 NUMERIC-MOVE.
024000      IF STRATA-SIGNAL IS EQUAL TO DATA-ZERO,
024100      GO TO DATA-ZERO-PRINT.
024200      IF STRATA-SIGNAL IS EQUAL TO G-STRATA,
024300      MOVE TEMPORAL TO G-AREA,
024500      GO TO STRATA-PROCESSING.
024600      IF STRATA-SIGNAL IS EQUAL TO P-STRATA,
024700      MOVE TEMPORAL TO P-AREA,
024900      GO TO STRATA-PROCESSING.
025000      IF STRATA-SIGNAL IS EQUAL TO S-STRATA,
025100      MOVE TEMPORAL TO S-AREA,
025300      GO TO STRATA-PROCESSING.
025400      IF STRATA-SIGNAL IS EQUAL TO T-STRATA,
025500      MOVE TEMPORAL TO T-AREA,
025700      GO TO STRATA-PROCESSING.
025800 SIGNAL-SET.
025900      IF LETTER IN WORK-AREA IS EQUAL TO G-STRATA,
026000      MOVE G-STRATA TO STRATA-SIGNAL,
026100      GO TO STRATA-PROCESSING.
026200      IF LETTER IN WORK-AREA IS EQUAL TO P-STRATA,
026300      MOVE P-STRATA TO STRATA-SIGNAL,
026400      GO TO STRATA-PROCESSING.
026500      IF LETTER IN WORK-AREA IS EQUAL TO S-STRATA,
026600      MOVE S-STRATA TO STRATA-SIGNAL,
026700      GO TO STRATA-PROCESSING.
026800      IF LETTER IN WORK-AREA IS EQUAL TO T-STRATA,
026900      MOVE T-STRATA TO STRATA-SIGNAL,
027000      GO TO STRATA-PROCESSING.
027100      IF LETTER IN WORK-AREA IS EQUAL TO DATA-ZERO,
027200      GO TO DATA-ZERO-PRINT.
027300 ERROR-ROUTINE.
027400      MOVE SPACES TO PRINT-AREA. MOVE O-BLOCK-NUMBER
027500      TO BLOCK-NUMBER IN PRINT-AREA.
027600      MOVE "STRATA FORMAT INVALID" TO WORD IN PRINT-AREA.
027700      WRITE PRINT-AREA BEFORE ADVANCING 2 LINES.
027800      GO TO READ-INPUT-FILE.
027900 ADD-ONE-TO-I-POSITION.
028000      ADD ONE TO I-POSITION.
```

で，本来なら両方ほしいところである。

このようなことについては，この報告論文の末尾に実例が示してあるので，参照してほしい。

以上述べたように，この COBOL-KWIC はいくつかのプログラム・グループから成り立っている，全一体のプログラムシステムである。次にそのひとつひとつについて，やや立ち入って説明をしてみたい。

3. 長単位処理ルーチン

ここでは長単位処理のチェック・プログラム・グループについて述べる。ここに含まれるプログラムは，LUP-0, LUP-01, LUP-02, LUP-1, LUP-10, LUP-2, ソート, LUP-3, ソート, LUP-4 などである。LUPは Long Unit Proccessor の意で，テープラベルのＬＵＤのＤは Data である。このプログラム群の目的はさん孔された紙テープのメカ的誤さん孔をチェックし，層別等についての人間作業をある程度チェックし，必要があればデータのメンテナンスを行ないえるようにし，一語一語のレコードを作成して文内度数を数えたファイルをつくる，という点にある。

次に一つ一つのプログラムについて，大体を説明することにする。くわしい

```
028100      IF I-POSITION IS GREATER THAN I-AREA-FULL,
028200      ADD ONE TO I-AREA,
028300      MOVE ONE TO I-POSITION.
028400 ADD-EXIT. EXIT.
028500 SUBTRACT-ONE-FROM-I-POSITION.
028600      SUBTRACT ONE FROM I-POSITION.
028700      IF I-POSITION IS EQUAL TO ZERO,
028800      SUBTRACT ONE FROM I-AREA,
028900      MOVE I-AREA-FULL TO I-POSITION.
029000 SUB-EXIT. EXIT.
029010 DATA-ZERO-PRINT.
029020      MOVE SPACES TO PRINT-AREA.
029030      MOVE O-BLOCK-NUMBER TO BLOCK-NUMBER IN PRINT-AREA.
029050      MOVE "DATA ZERO" TO MESSAGE IN PRINT-AREA.
029060      WRITE PRINT-AREA BEFORE ADVANCING 2 LINES.
029070      GO TO READ-INPUT-FILE.
029100 RERUN-ROUTINE.
029200      CLOSE SPECIAL-FILE REEL.
029300      GO TO READ-INPUT-FILE.
029400 END-ROUTINE.
029500      CLOSE INPUT-FILE, OUTPUT-FILE, SPECIAL-FILE,
029600      PRINT-FILE. STOP RUN.
```

ことは，プログラムのなかを読まなければわからないが，くりかえしていうようにプログラムはCOBOLで書かれており，その書き方も人に読まれることを予想してできるだけわかりやすいように工夫してあるつもりであるから，オペレーションなどの上で疑問を生じたら，一度読んでみることを希望する。

LUP-0.（L 0と略称，以下同じ）

これは紙テープデータをよみこんで最後に end of item 記号を書き入れたファイルをつくることを目的としたプログラムである。紙テープデータのエンドマークとして，どの紙テープのうしろにも=E/Bのマークがはいる約束になっている。このプログラムではこれを読むと自動的にリランをとるようになっているので，紙テープの終わりでいつでもしごとがやめられるようになっている。紙テープのはじめのブロック番号のところだけがラインプリンタで印字される。もし出力ファイルを閉じたければ CLOSE INPUT-FILE. というデータを入れればよい。

このプログラムは FCP を使用したアセンブリ言語で書かれている。これはCOBOL で READ 命令を発すると，あらかじめ COBOL FORMAT に編集した紙テープしか受けつけないため，COBOL が使いにくいからである。もしどうしてもCOBOLを使うのであれば ENTER を使うよりほかないだろう。ここではそうしなかった。

LUP-01.（L01） このプログラムは LUP-0. でよみこんだデータ・ファイルが複数巻に分かれているものを一本にまとめるプログラムである。プログラムはベーシック・アセンブラで書かれている。

このプログラム，あるいは LUP-10 のいずれかによってファイルをまとめればよいが，後者は処理速度がおそい。

LUP-1.（L1）

このプログラムのリマークの項に，

REMARKS. THIS PROGRAM CHECKS MECHANICAL PERFORATION ERRORS OF KANZI-TELETYPEWRITER.

とあるが，これがこのプログラムの目的である。このプログラムのなかでは，ブロック番号が正確にできているか，12ビットすなわち2ディジット1漢テレ

文字の区切れが最後まで完全か（どこまで1ディジットのぬけがないか），12ビット2ディジットのうちどちらかがスペースになることが多いが，どこでそれが起きているか，などをチェックした。以前のJ字チェック（参照報告31，143ページ）を省略し，かわりに右か左いずれか一方のみがスペースになるものを拾い出すことに改めた。そしてテープのなかのどの部分でエラーが起きているかを見やすく印字できるプログラム（紙テープにさん孔してある）を作成しこれを併用することにした。これはエラーの摘出について有効であり，それを用いて紙テープの修正を行なうばあいにも便宜が与えられるようになり，作業上のスピードが大幅に改善された。

木村繁氏が行なった LAST NUMBER-SIGN CHECK（報告31，143ページ参照）は有益有効なので，このプログラムでもこれをなぞった。これは原文がメカ的な事情で二つに切れてしまったばあいに威力を発揮する。処理したものについて，ブロック番号のうしろに，完全もなのについてはCOMPLETEと，欠陥のあるばあいはその種類と場所が印字される。

このプログラムのなかでは原文のなかに自動的にはいった漢テレの自動復改記号がとりのぞかれるが，ブロック番号を除き，復改コードを除いて，4000漢テレ字のものまで COBOL-KWIC では処理できるようになっている。したがって LUP-0 で処理する紙テープでは漢テレ4150字（HITAC3010では8300字程度まで）の入力は可能である。しかしそれをこえると “AREA OVER” の印字が出る。

このうしろに LUP-10. というプログラムも作成してある。これは LUP-1. で処理したいくつかのファイルを合わせて1つのファイルにするプログラムである。

LUP-2.（L2）

このプログラムは内容を示した。そこに REMARKS があるが，別に新聞用語調査のデータでなくともよいので，とにかくスペースで区切られたところまでの文字連続を一語として取り扱い，これを1レコードにするプログラムである。レコードは，ブロック番号，文番号，単語番号，層および見出し語のようになっている。

これらの文や語のカウントをなかでやっている。

このプログラムのなかで，原文のなかにあった層に関する情報が編集されて各語につけられる。GPST（参照，報告31，3ページ）のそれぞれの数値がはいる。新聞用語調査以外のばあい，GPSTに勝手な意味と勝手な数字を与えることがゆるされ（数値は二けた以下），何も与えないことも許される。ただ層の表示法で欠陥があるばあい，たとえばGPST以外の文字が現われたばあいとか，層表示終了のマーク“）”がないばあいは，警告を発し（STRATA FORMAT INVALID），そのデータはネグレクトされる。

もとの新聞のデータでブロック内に文字がなかったばあい，も印字される“DATA ZERO”。

ソート1。ブロック番号，センテンス番号，単語番号。このソートは次のLUP-3．のために行なう。

LUP-3．（L3）

REMARKS. THIS PROGRAM HAS TWO AIMS, ONE OF THEM
IS TO CHECK THE LONG UNIT INPUT DATA FORMAT
ESPECIALLY CONCERNING STRATA INFORMATION,
THE OTHER IS THE MAINTENANCE OF DATA
ESPECIALLY DOUBLED DATA.

これは，デリートしたいデータのデリート，2重以上にはいったデータを一通りだけ残すこと，GPSTのうち0になるものかあるいは一定以上の数値のもの，たとえばGなら18以上のもの，はフォーマット作成上の（人間の）エラーとみとめて，そのデータを取り除く。

新聞語い調査以外のデータはこのプログラムを通さない方が安全である。通しても通さなくても，磁気テープ上の記録は変わらない（ラベルだけは変わる）。

データをデリートするときは，ＩＮＴボタンを押して紙テープをよみこませてからメインプログラムにうつる。デリートデータはブロック番号で7けたとe/i，これをノン・ラベルのフォーマットで紙テープさん孔する。

ソート2。見出し語漢テレ順，文番号，単語番号順。次のLUP-4．にそな

える。これはまた KWIC の TABLE-COLLATE の入力ファイルでもある。

LUP-4．(L4)

REMARKS. FREQUENCY COUNTING OF WORD
IN A SAME SENTENCE.

どの単語でも現われた限り少なくとも1回とカウントされるわけだが，このプログラムのなかでは同一文内での出現度数をカウントするだけである。これは本来それほど必要なステップでもないが，すでにできあがっている長単位処理プログラムではこうなっていないと動かないようになっているので，これを入れたのである。単語レコードを作成するステップのなかでやってもよいことだが，KWICにつなげるためにはこのようにした方がよく，文内ソートをするよりもこのようなプログラムをつくった方が時間が経済だからである。このプログラムによって，磁気テープファイルの数を多少少なくすることができた。

このプログラムでは BLOCK-NUMBER と SENTENCE-NUMBER をまとめて NUMBER-INFORMATION と定義しているが，このプログラムの中心は

IF WORD IN INPUT-AREA IS EQUAL TO WORD
IN LAST-DATA AND NUMBER-INFORMATION
IN INPUT-AREA IS EQUAL TO NUMBER-INFORMATION
IN LAST-DATA, GO TO FREQUENCY-COUNT,
OTHERWISE GO TO WRITE-OUTPUT-FILE.

のように表現されている。

このプログラム終了後，データは，長単位処理後半部分，RUN1 以降（斎藤秀紀氏，花井夕起子嬢作成のプログラム）へと渡されるのである。

4. 漢字解読用例印字ルーチン

このプログラムグループは，長単位処理の途中からデータを受けてこれを KWIC 印字へとつなぐ役目をするが，インプットが漢字かなまじり文，アウトプットがかな文であるため，その漢字部分をかなに改める役目をする。ふつ

うの形の日本文を読みくだすという作業をするわけである。

ここで二つほど取りあげたい問題がある。それは漢字解読のシステム・アナリシスとそのなかでの辞書検索の方法についてである。

漢字解読のシステムとしてはすでに田中章夫氏のものがある（国語研報告34，107ページ）。これは一定の小さなテーブルを用いて短時間に処理できるが，このテーブルがはん用性をもっている点は特に便利である。ただ現在のところ固有名詞，連濁・連声・音の交替・あて字・熟字訓などは，本格的な処理を目ざしていない。

新聞用語調査のばあいは，漢字にかなをつけたファイルが，すでにできている。かなふりは長単位の語について人間が行なったものなので，その範囲内ではかなり正確なはずである。これを用いれば，連濁・連声・音の交替・あて字・熟字訓などは正確に処理できる。「一〔いち〕般〔ぱん〕に夏〔なつ〕すぎから上〔じょう〕向〔こう〕くという観測が有〔ゆう〕力〔りょく〕なようだが」，のようなばあい，漢字がはいっているから読みやすいが，漢字をぬいて「イチパン　ニ　ナツスギ　カラ　ジョウコウク　ト　イウ　カンソク　ガ　ユウリョク　ナ　ヨウダガ」となるとわかりにくい。「年〔とし〕が明〔あ〕けても引〔いん〕続〔ぞく〕き行〔ゆ〕われ」（以上2例，（報告34，128，127ページ）を「トシ　ガ　アケ　テモ　インゾクキ　ユワレ」となると，読みにくい。このような純粋かな文では，連濁なども正しくできていた方がよみやすい。また新聞では人名・地名など固有名詞が非常にしばしば出てくるので，この点も解決したい。そのためには，新聞用語調査の用例印字の範囲では，上述の辞書を使った方が，より有効であろう。ただ問題点としては後にくわしく述べるように，二様以上に読み分ける必要のある語表記が読み分けられないのだが，これは語単位なのでさいわいにしてそうひどく多くはない。事実実際にプログラムを作成し検査したところ，漢字の正読率は，固有名詞，連濁・連声などをすべて計算に入れてもなおかつ99パーセントを上回る好結果となった。これなら実用プログラムとして使用できると思われたので，さっそく使用したのである。このような文字よりも長い単位で扱う方が解読率がよくなることはすでに田中氏も言及されている（報告34，137ページ）。

このように，文字単位よりも語単位の方がよい結果が出るのはどうしてだろうか。

これは，日本語のなかでの漢字使用が，通常は語の表記にあてられているからだ，と考えることができるのではあるまいか。語といったが，形態素といってもよい。アラワレルを「現われる」とも「現れる」とも書くことができるのは，そのためであろう。タバコを「煙草」，ウルサイを「五月蠅い」と書くことができるのも，その一つのあらわれにすぎないのではないか。ウが五にあたるのではないから，このばあいのウルサイと五月蠅いとは，シドニー・M・ラムの stratificational grammar の表示法では，多分少なくとも一つ上の層を通してしか結びつかないであろう。それはすなわち語あるいは形態素を表記するということにつながるのではないか。実は日本語では，漢字もかなも，ともに語を表記するためにあるといえる（中国語の事情はわからないが，おそらく語を表記するためにあるといえよう)。

語を表記するのであれば，漢字を読むばあいに語として読むということは，当然の方法である。この漢字解読はこのレベルで行なっているので，結果がいいのも当然であろう。ただし同じ文字連続（一字のも含めて）において二通り以上の読みのあるもの，たとえば「高潮」の「たかしお」「こうちょう」，「国立」の「こくりつ」「くにたち」，「今日」の「きょう」「こんにち」，「左側」の「さそく」「ひだりがわ」，「夫婦」の「めおと」「ふうふ」，「工夫」の「くふう」「こうふ」（参照LDP 6）などや，「間」の「けん」「かん」「はざま」「あいだ」「ま」などは読みわけができず，どれか一つがはいるようになっている（はいるものがいつでも一つにきまっている)。これらについていえば，文のなかにおいて語がきまるのであって，やはり語の認定ができないから読みがきまらないのにほかならない。この種の二通り以上あるものは，そのいずれか一つがはいっているので，このばあいには誤りになるわけであるが，これらの語が比較的少ないので，そのままにしておき，その判定を文脈を人間がみて定めうるようにしたのである。

辞書を用いるとすると，辞書に出ていない語が出てきたらどうするかということを考えなければならない。このばあい，手軽には人間が教えるのがよい

し，はじめの趣旨にもかなうだろう。そこで，辞書にない用語をリストして人間に教示をこわせるように設計した。これを教えてやれば，その単語についての読みを理解し，辞書をふやすことができよう。

辞書を作成・保守するルーチンは，別につくった。これについては，6．の付属ルーチンの項で述べる。

さて，辞書をひくという設計のばあい，最も問題なのはテーブル・ルック・アップのプログラム手法である。これをへたにつくると，小さなテーブルを引くのにも意外な時間がかかる。磁気テープをいちいちリワインドしたりすると手間だけでなく，ファイルの損傷も考えなければならない。大量の処理のばあいにこれを気にしているようでは困る。文字でなく語単位で扱うばあいの問題点が実はここにあるので，プログラム的に解決してゆかねばならない。この点について，筆者はさきに，計算機のなかでのインナー・ソートを行ない，それからテーブル・ルック・アップを行なうことを考えてみた。これを実験してみて，大変効果があった。このことは小文「電子計算機による語い調査の一実験」『ことばの研究』2で報告ずみである。これは，ハーバード大やソ連での実験でもその例があるらしい（G. Mounin : La machine à traduire. 1964 Mouton)。 筆者のさきの実験のばあいは筆者が作成したソートプログラムを利用して，計算機本体のなかで一文ずつソートを行ない，一文ずつテーブル・ルック・アップを行なったのであった。この COBOL-KWIC のは，二通り以上の読みがあるものについて読み分けを行なわないことにしたので，一文ずつのテーブル・ルック・アップでなくてもよい。すなわち文でなく全文章一挙のテーブル・ルック・アップであってよいのである。思い切ってこのような設計にしたところ，非常に早くテーブル・ルック・アップができるようになった。すなわち，辞書10万語（異なり），文章10万語（10万α＝15万β，延べ）でテーブル・ルック・アップは40分でできた。

このようなコレート式テーブル・ルック・アップでは前とうしろにソートが心要であるが，さいわいにして，前のはすでに長単位処理のにふくまれているので，すんでしまっている。これをそのまま使用すればよい。うしろのはブロック番号などによってもとの文章の順にもどすものであり，これは行なわなけ

ればならない。このソートは，サービス・ルーチンとして提供されたものを使用した。

この種の辞書引きのルーチンは，漢字の解読ばかりでなく，その他の作業のばあいにも有効である。語種や品詞など，いわゆる付加情報の付加処理のばあいにも有益な手法であろうと思う。その方法は table collate 自体のプログラムの修正なしで，辞書を変えるだけで可能である。

そこで，この COBOL-KWIC のメイン仕様では，漢字を解読するだけでなく，長単位を短単位に切る作業も同時に行なわせたのである。これは辞書がそのようになっていればよいだけで，メイン・プログラムにはいささかの変更も要しない。だから辞書をかえれば，それだけで短単位に切らないようにすることもできる。一応短単位にしたのは，その方が KWIC として役に立つものができるからである。

次にこのルーチンを構成する，各プログラムについて概略説明する。

TABLE-COLLATE. (COL)

長単位処理 LUP-3．のあとのソートのすんだデータをインプットとし（これは漢テレ順に並んでいる），辞書（これも漢テレ順にならんでいる）と対照して漢字の読みと短単位のくぎり目のついた情報をそのまま書き写して，ブロック番号，文，単語番号に添えたデータをアウトプットする。辞書にない見出し語は紙テープに出す。このプログラムの基幹部分は，インプット・ファイル，テーブルファイルを読んだあとの次の部分である。

```
IF WORD IN INPUT-FILE IS EQUAL TO
HEAD-WORD IN TABLE-FILE, GO TO WRITE-OUTPUT-FILE,
OTHERWISE IF GREATER, GO TO READ TABEL-FILE,
OTHERWISE IF LESS, GO TO WRITE-UNMATCHED-FILE.
```

漢テレ順に並んだものどうしのコレートであるから，TAPE SWAP 以外のリワインドはない。

ソート3。ブロック番号，文番号，単語番号による。

WORD-TO-SENTENCE-MAKING. (WS)

単語ファイルから文ファイルに改めるプログラムである。ブロック番号と文

番号をあわせて NUMBER-INFORMATION とし，次のように書いたステートメントがこのプログラムの眼目である。COMPARE-AREA には一つ前のデータがはいっている。

IF NUNBER-INFORMATION IN INPUT-AREA IS EQUAL TO
NUMBER-INFORMATION IN COMPARE AREA,
GO TO SENTENCE-GATHERING.
GO TO WRITE-OUTPUT-FILE.

なおこのプログラムは新聞用語調査の用例印字のばあいには漢字解読ルーチンの一環として使用されるが，このプログラムのごく一部を変えた WORD-TO-SENTENCE-MAKE-K (WSK) というのも用意した。これは最初から全部かなで入力したもの，あるいは「解〔かい〕説〔せつ〕」のように最初から漢字にかなをつけて入力したものを取りあつかうことができるようにしたものである。このプログラムのインプットデータは LUP-2. の処理を受けたものすなわち "LUD-2" である。

KANA-TEXT-MAKING. (KTM)

上の WORD-TO-SENTENCE-MAKING. あるいは WORD-TO-SENTENCE-MAKE-K. のアウトプットをインプットし，漢テレコードのかな文字による純かな文作成を行なうプログラムである。インプットの「解〔かい〕説〔せつ〕」を，「かいせつ」に改める。

このプログラムは現在そのようになっているが，これに手を加えれば，「解〔かい〕説〔せつ〕」をたとえば「かい〔解〕せつ〔説〕」に直すことも考えられる。このようにしてできた KWIC は表記の状態が調査できるようになるだろう。

この KANA-TEXT-MAKING. のプログラムによって，漢字解読ルーチンは終了する。

5. 用例印字ルーチン

前述の漢字解読で作成した完全なかな文を，ラインプリンタで印字するまで

の諸手続きを形成するプログラム・グループである。主なしごととすると，漢テレのかなコードをライン・プリンタのコードに直すこと，原文の一覧表をつくること，KWICのフォーマットをつくること，それを印字することである。

CHARACTER-CONVERSION. (CC)

前節でできた完全かな文について，ラインプリンタで印字ができるようにかななどの漢テレコードをラインプリンタのそれに直す方法である。このプログラムでは，コンバージョンをテーブル・ルック・アップの方法で行なっているので，少し手直しすればかな文字でなく，ローマ字で印字するようにすることもできよう。このように，主としてテーブルの操作によってプログラムの基本的な点に手を入れなくても他の種の変換ができるようになることは大変いいことであるが，原文の１字１字をとりあげ，これをテーブルのなかの１字１字と照らしあわせながらコンバージョンするというのは，時間的にみてやや長くかかる欠点がある。本来なら辞書の段階で直しておくのがよいだろうし，あるいは文字ごとのレコードをつくり，ソートするという方法もあるかもしれない。それをしなかったのは，あくまで漢字情報入りの KWIC が必要だと考えていたからである。

COBOL でのこの種のデータの取り扱いは，プログラムを書くばあいにはきわめて便利であるが，作成したプログラムはラン・タイムではかなりの時間を要することになった。しかし ENTER 動詞を使用しなかった。

漢テレの文字数とラインプリンタの文字数では，漢字をぬきにしてもかなりの開きがある。特に記号類，小さいッャュョィェォ，「を」などはやむをえない。小さいッャュョ類は大きい字で代用し，足りない記号類はすべて“＋”で代用，「を」は最後に並ぶようにしてオと印字（64文字で＝＊）することにした。

TEXT-PRINT-OUT. (TPO)

ライン・プリンタ・コードになったかな文字を編集して用例一覧表をつくるプログラム。このプログラムの中心は

WRITE PRINT-AREA, BEFORE ADVANCING 2 LINES.

にある。前にも述べたが，KWIC の用例印字が前後一定字数のところで切れ

ているので，この印字があるともとの文にもどれて便利である。これはブロック番号と文番号で検索することが可能である。

KWIC-FORMAT-MAKING. (KFM)

もとの文の各単語をとりあげて見出しとし，その前後の用例を編集して，次のプログラム KWIC-PRINT-OUT の入力ファイルを作成する。同時にソートキーを指定しやすい形につくる。このプログラムの論理は別にブロックチャートで示したが，ENTER 動詞を使用せず，純粋に COBOL で書くと便利ではあるが実際はやや複雑になる。このためランタイムで時間がかかることになった。

ソート４。見出し語，直前語あるいは直後語，直後語あるいは直前語をソート・キーとして分類する。直前直後の相違についてはすでに述べた。

KWIC-PRINT-OUT. (KPO)

最終的なプリント・アウトのプログラムである。

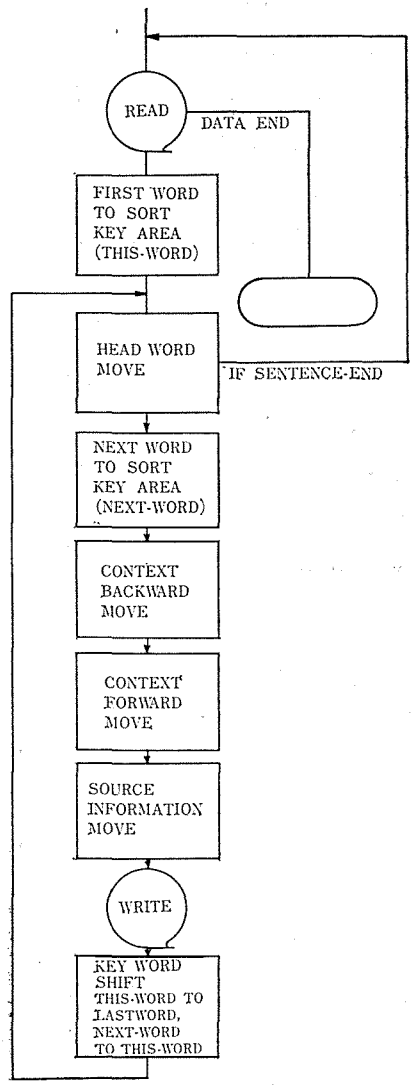
6. 付属ルーチン

辞書の作成と修訂のルーチン。

KANZI-BETA-TABLE-MAKING. (KBT) 前に述べたように，用語調査短単位処理プログラムのなかで作成される “BETA-011” というファイルから，そのβ３にあたる，漢字のよみと短単位切りの情報をもった部分をぬき出して，TABLE-COLLATE で使う辞書 “KB-TABLE” をつくるプログラムである。このプログラムを改造してβ２を取り出すとかなだけのものができる。辞書は固定長にしておいたが，これは辞書が大きくなるマイナスより，検索時間が長くなるマイナスを防ぐようにしたためである。逆のばあいの実験をしてみなかったが，固定長にしたことは，有益だったと考えている。

MAINTENANCE-TAPE-READ. (MTR) TABLE-COLLATE で UNMATED-FILE として出てきた語のリストにつき，人間が漢字のよみ方を指示した内容をさん孔した紙テープを読みこみ，辞書のフォーマットを作成するプログラム。

KWIC-FORMAT-MAKING.
BLOCK-CHART.



DICTIONARY-MAINTENANCE. (DCM) 上述の読みこみ内容をトランザクション・ファイルとして，マスターファイルを更新するプログラム。指定されたフォーマットにしたがって削除，追加，訂正ができるようになっている。

用例検索 (IR) ルーチン。

SPECIFIED-DATA-EXTRACT. (SDE)

研究者が欲する語の用例だけを求めるばあいに使用するルーチンである。紙テープに欲する語を 3010 の 333C カナコードで入れるとリクエスト・ワードだけのファイルができる。これを前または後の語を優先してソートするか，またはしないで， KWIC-PRINT-OUT のルーチンにかければ，その用語だけの KWIC ができるようになっている。 紙テープは96文字コードで語を書き，全部で20けたになるようにあとをスペースでうめ， e/i をつける。全体はノンラベルの紙テープフォーマットにする。

7. 問題点など

このプログラムの，長単位処理の部分を使ってすでに200万語が処理された。大部分は問題なく終わっているといってよいのであるが，一部に連絡不十分のための事故が出ている。それは，パンチの上でおしてはならない盤外キーを誤まって押してしまったばあい，長単位のプログラムの後の方で故障がおきたのであった。ただこれはきわめて例が少なく，かつきまったタイプのエラーなので，特別のプログラムをつくって長単位処理を終わらせてしまった。この事故は COBOL-KWIC 自身のなかでは少しも影響をこうむらないように，安全装置がついていたので，COBOL-KWIC のなかの長単位処理プログラムはいじらないことにしているのである。

次に辞書の作成時に，辞書があまり大きくなることをおそれて情報を漢テレ40字でおさえてしまったことも問題である。そのために長いものを途中で切ることをあえてした。ただできあがった KWIC をみていると，この種の例はきわめて少ないために，ほとんど気がつかない程度である。もし将来このプログラムに手を入れるとしたら，この点が問題になろう。

このプログラムを実際に全部通して運転してみた感じでは，問題点は CHARACTER-CONVERSION と KWIC-FORMAT-MAKING にある。この二つが全体の所要時間の五分の三以上を占めるのである。それは主として中央処理装置のなかでのしごとなので，全体の効率は中央処理装置の速度の向上によって，かなり改善されよう。10万α語（＝15万β語）での全処理で約百時間かかったが，さきの二つのプログラムのほかで時間がかかるものは，紙テープの読みこみという人間作業を伴うものでしかもメカ的な限界があるもの，およびソートであって，これは作業の性質上いかんともしがたい。

このプログラム作成の目的ははじめに述べたが，これに関して言い添えておきたいことがある。このプログラムが主として新聞用語調査の用例印字によって同語異語の判定に役立つことを考えたこと，および，語い文法表記文体など言語学的な各種の分析に必要なデータを供給するためであることはすでに述べたが，このそれぞれについてである。

まず同語異語の件であるが，同語異語の操作は単位切りとならんで，用語調査の基本である計量単位の基本点であるということができよう。新聞用語調査では単位分割の方は一応形をととのえているが，同語異語については全く手をつけていないといってよい。これは明らかな片手落ちであって，このままでは計量単位として半分はできているが，半面はゼロに近いという不完全な語い表になってしまった。この新聞用語の調査自身の欠を補う必要のあることは当然であるが，次の用語調査を設計するに際しても，このことは現段階の最大の課題としてとりあげなければならないはずである。その自動化ができるかどうか，筆者としてもプログラムによる実験をしたことはないではないが（論集2），当面すでに，全面的に行なうことはむずかしいかもしれない。だとすれば不完全な自動化はその修正の手間と，エラーの見のがしなどの点から，むしろ，次の用語調査はむしろ KWIC を主体にするという考え方を持ってもよいのではないかと思う。このばあい，できれば漢字の高速プリンタを備えて原表記のままの KWIC ができることが理想的であろう。これはメカ的なコンディションによるところが大きいのであるが，この COBOL-KWIC はそのような点まで見通して，設計したところがある。すなわち，そのようなプリンタが入

手できたばあいでもなお有義をもつであろうものを，最優先で作成したのである。そのプログラム手法は機種が変更されてもそのまま生かされるであろう。

フランスで行なっている世界一の規模をもつ大用語調査についての情報は，いまだ十分に入手していないが，先年日本で行なわれた FID の大会で，フランスの F. Lévy 氏から聞いた話では，この作業の中心は “faire des concordances” であるということのようであった。用語調査の基本は多分この面でも concordances をつくることであろうと思う。この点からも（KWICのフォーマットでなくてもよいが）用語調査は KWIC 的なものを主体にすべきであると考えられる。

用例印字の結果が文法・語いなどの言語学的研究に有意義であることはすでに述べたが，そのような文法・語い論的な研究が逆に言語情報処理に寄与することが多いことはすでにのべた。このようなデータは電子計算機による日本語の基礎的研究にも有益である。用語調査を計算機で行なうことは言語情報処理の一つと考えるべきであるから，語い調査の結果が語い調査にかえってくるということも十分考えておかなければならない。このようにして用語調査の方法が少しずつ改善されなければならないのである。そのような意味でさきに作成した自動分かち書きのプログラムの文法・語い論的基礎としてこの COBOL-KWIC で作成したデータが十分に利用できると思われる。このことも当初からねらった点であった。

以上のような理由から，このプログラム作成に続く問題として，次の二つのことが出てくる。一つは将来漢字の高速プリンタが導入されたばあい，このプログラムの手法のうち欠点を改め，よい点を採用して，用語調査の重要部分をなすものとして，新しい機種に即したプログラムの条件を考えてみることである。COBOL で書いてあるということが，この際重要な意味をもってくるわけである。もう一つは，このプログラム自身のアウトプットを利用して，言語学的な調査研究を行なうことである。例に示した助詞ヲ＋動詞モツの目的語として，どんな名詞がくるかというような分析は，言語情報処理にとって有益な基礎資料となるであろう。

なおこのほかに，このプログラ作成の動機として，実務として長単位処理を

なんとかしなければならないという必要があったこと，プログラム言語としてコンパイラ使用を検討してみる（世間の常識からすればやや遅きにすぎるが）ことなど，いくつかあるが，すでに言及してあるので，ここではくりかえさない。

このプログラムは新聞用語調査の入力フォーマットに大体従えば，多少のずれがあっても扱うことができるようになっている。もしはじめからすべてかなで入れたとすれば，このなかの漢字解読ルーチンをとばして先へ進めればよい。層の表示も新聞以外のものでは，あまるものと不足のものとあろう。ＧＰＳＴを勝手に定義すれば，自分なりの使用法ができる。２けた以内の数字なら自由に使ってよい。ただしこのようなばあいには，LUP-２．のあと漢テレ順のソートをしていきなり WORD-TO-SENTENCE-MAKE-K. に連続させなければならない。そのあと KANA-TEXT-MAKING. 以下を通すのである。このばあい KAKA-TEXT-MAKING. を通すのは，不要なようであるが，実はそうすることによって，「学〔がっ〕校〔こう〕　へ　行〔い〕く」のような入力紙テープも同様に処理できるようにしたためである。短絡のプログラムをいちいち作らないで，類似のものはなるべく道を一つにしてプログラムの本数をへらしたのである。

ここで作成した KWIC は KWIC のひとつの見本にすぎない。他の種の KWIC が考えられる。意味の番号や品詞などくわしく記入した KWIC は，さらに有益であろう。さきにも述べたように言語研究の最も基本的なデータであると思われるので，さまざまな形の KWIC が出ることを期待する。特に単語の意味情報をつけた KWIC は文法，語い論的研究にとって有益であると思われる。もしこの COBOL-KWIC を延長するとすれば，辞書の形を工夫することによって解決できよう。

このプログラムの作成は大体筆者が行なったが，すでにしるしたように，データマージの段階でのプログラムを斎藤秀紀氏にお願いし，キー・パンチャーが inhibit character を押したばあいのあとしまつのプログラムを花井夕起子嬢にお願いした。オペレーションとプログラム修正の一部を花井嬢にお願いした。これには組みこまれていないが，より充実した検索システムの作成を村木

プログラム・オペレーション一覧

	入　力	出　力	その他	リラン (ふつう INT)	outputMTの ラベル	言　語
LUP-0	PT	1R(注)	①	自動,2R	ノン・ラベル	アセンブラ
LUP-01					ノン・ラベル	アセンブラ
LUP-1	1	3R, LP		不可,2R,4R	LUD-1	COBOL
LUP-10	1,2	3R		不可	LUD-1	COBOL
LUP-2	1	3R, LP		可,2R,4R	LUD-2	COBOL
LUP-3	1(PT)	3R, LP	②	可,2R,4R	LUD-3	COBOL
LUP-4	1	3R		可,2R,4R	LUD-4	COBOL
COL	1	3R, LP	5(KB-TABEL)	可,2R,4R	KWICTAPE	COBOL
WS	1	3R		可,2R,4R	KWICTAPE	COBOL
WSK	1	3R		可,2R,4R	KWICTAPE	COBOL
KTM	1	3R, LP		可,2R,4R	KWICTAPE	COBOL
CC	1	3R		可,2R,4R	KWICTAPE	COBOL
TPO	1	3R, LP		可,2R,4R	KWICTAPE	COBOL
KFM	1	3R		可,2R,4R	KWICTAPE	COBOL
KPO	1	LP		可,2R,4R		COBOL
KBT	1	3R		可,2R,4R	KB-TABLE	COBOL
MTR	PT	3R, LP		可,2R,4R	KB-TABLE	COBOL
DCM	1	3R		不可	KB-TABLE	COBOL
SDE	1	3R	③	不可	KWICTAPE	COBOL

（注）Rはライト・ロックアットリングをかけること。数字はテープ・デッキ。

注意

①紙テープの終りに＝E\B があるとリランをとる。

②はじめに INT を押すと DELETE-DATA の BLOCK-NUMBER をよみこむ。

③INT をおすとしごとを中止する。

氏にお願いした。それぞれ，謝意を表する。

8. フォーマットおよびデータ見本

フォーマットについて。

LUD-2．から LUD-3．のデータフォーマットは共通でこれはプログラム（LUP-2）に示されている。LUD-4．のフォーマットは単語番号の部分がカウントになっているだけであとは変わらない。

辞書のフォーマットは，BLOCK CONTAINS 5 RECORD で1 レコードは，

```
01  TABLE-AREA.
    02  HEAD-WORD, PICTURE IS X (40).
    02  INFORMATION, PICTURE IS X (82).
```

である。 TABLE-COLLATE のなかで HEAD-WORD が INPUT と比較されるところで，たとえば「東京……」のようになっている，INFORMATION には「東〔とう〕京〔きょう〕……」のような形でデータがはいっている。だから HEAD-WORD が一致したら，INFORMATION を MOVE するだけでよい。

TABLE-COLLATE の UNMATCHED-FILEは，辞書になかった用語がリストされ，漢テレ順に並んでいる。これらの単語の読み方（あるいは短単位の切り方）を教えれば，次回からは既知語として処理されるようになる。そのためには，次のようなフォーマットを作成して MAINTENANCE-TAPE-READ, にかければよい。一レコードは HEAD-WORD C/R INFORMATION / C/R C/R e/i ギャップとする。HEAD-WORD は variable length で漢テレ20字以内，C/R は漢テレ改行復帰，／は漢テレスラントで e/i は3010の記号のであり，INFORMATION は辞書の種類によって異なるが，“KB-TABLE” のばあいは漢テレ40字以内の variable length で，たとえば「学〔がっ〕校〔こう〕 教〔きょう〕育〔いく〕」のようにする。このようなフォーマットで，non standard label の形式の紙テープをつくって入れればよい。

データおよび印字見本についての説明

① 最初のインプット・データを漢テレで印字したもの。毎日新聞夕刊の社会面ニュースの例。＃は文の切れ目のしるしに入れてある。

② 漢字解読したあとのデータ（WORD-TO-SENTENCE-MAKING）を特に編集して漢テレで印字したもの。読みがなは漢字のうしろの〔 〕のなかにはいっている。データのはじめの数字はブロック番号7けた，層情報8

けた，文番号2けたを示す。

③ 上の②のかなを拾って完全なかな文にし，ラインプリンタで打ち出したもの。すなわち TEXT-PRINT-OUT, のアウトプットである。

④ 上のテキストについて各単語を見出し語として前後の用例を添えたもの。すなわち KWIC-FORMAT-MAKING. のアウトプットで，これを特にこのまま印字したものである。1行が1レコードになっている。もとの文の順序でレコードがつくられるため，全体としてみると1語ずつななめにずれているようにみえる。

⑤ KWIC の一般的な例。中央たてに通ったスペースの右側が見出し語である。その左右がコンテクストであるから用語がコンテクストのなかに置かれていることになる。左側の数字はブロック番号，文番号，層情報（GPST）である。＋は特殊記号があることを示している。－は引用符号「　」の代用。

⑥ KWIC の例の一つで，「的」がどんな語につくかがわかるようになっている。テキの直前の語が五十音順に整理されて並んでいる。もし後ろ優先のソートをすれば「テキ　ダ」「テキ　デ」「テキ　ナ」「テキ　ニ」などがきれいに整理される。このように KWIC は語構成の研究にも役立つ。

⑦ KWIC の例の一つで，助詞「を」の例。プリンタに「ヲ」がないので，「オ」で代用したが，「ヲ」が集まるような工夫がされている。これは後ろ優先のソートなので，助詞のうしろにくる動詞などが五十音順に整理される。したがって「を」をとる動詞にどんなものがあるかが一覧できるようになっている。ここに示した例は動詞「モツ」のページの一例で，助詞「を」の前の語が五十音順に整理されているので，連語「……をもつ」がどんな目的語をとるかが一覧できるように計算機が整理してくれている。このように，KWIC 語の意味用法はもちろん，文法，シンタクスの研究にも非常に有益な資料を提供する。ここにはテキとヲの例を示したがたとえばテの前を整理するとどんな動詞が連用形で使われるかがわかる。このように，形態論的な研究その他にも使用することができる。

①

（K001129）（G1P1S1T5） ＃ 大谷重工 深川工場 ボヤ （P4） ＃ 三日 午後 一時 十五分 ごろ，東京 墨田区 柳原町 一〇 ，大谷重工 深川工場 の 機械工場 南側 の カベ から 煙 が 出 て いる のを 従業員 が みつけ た 。 ＃ およそ 十平方メートル の カベ を こがし た だけで 消しとめ た 。 （P1） ＃ 埋立地 に 男 の 変死体 （P4） ＃ 〔 船橋 〕 三日 午前 三時 十分 ごろ ， 千葉県 船橋市 日の出町 一〇 ， 第一コンクリート会社 正門前 埋立地 の 臨海道路ぎわ に 若い 男 が うつ伏せ に なっ て 死ん で いる のを 通行人 が みつけ 船橋署 に 届け出 た 。 ＃ 新島は 品川区 大崎郵便局 の 集配人 と し て 勤務中 ， 一昨年 一月 十二日 ， 品川区 二葉町 を 配達 途中 で 千円入り 現金書留 一通 を 盗ん だのを はじめ ， 同年 十一月 までに 現金書留 六通 九千五百円 を 配達 の 途中 で 盗ん で 起訴さ れた 。 ＃ 一昨年 十二月 十八日 ， 一審 の 東京地裁 は 被告 に 徴役 一年 六月 ， 執行猶予 四年 を いい渡し たのに 対し ， 検察側 は 「 執行猶予 は 刑 が 軽すぎる 」 と 主張し て 控訴 ， 昨年 四月 十二日 ， 二審 の 東京高裁 は 検察側 の 主張 を 認め ， 一審 の 執行猶予 の 判決 を 破棄， あらためて 懲役 一年 の 実刑 を いい渡し た ため ， 被告 は これ を 不服 と し て 上告し て い た 。 ＃

②

K00112901010150 1 ＃ 大〔おお〕谷〔たに〕 重〔じゅう〕工〔こう〕 深〔ふか〕川〔がわ〕 工〔こう〕場〔じょう〕 ボヤ ＃

K00112901040105 02 ＃ 三〔みっ〕日〔か〕 午〔ご〕後〔ご〕 一〔いち〕 時〔じ〕 十〔じゅう〕 五〔ご〕 分〔ふん〕 ごろ ， 東〔とう〕京〔きょう〕 墨〔すみ〕田〔だ〕 区〔く〕 柳〔やな〕原〔はら〕 町〔ちょう〕 一〔いち〕 〇〔れい〕 ， 大〔おお〕谷〔たに〕 重〔じゅう〕工〔こう〕 深〔ふか〕川〔がわ〕 工〔こう〕場〔じょう〕 の 機〔き〕械〔かい〕 工〔こう〕場〔じょう〕 南〔みなみ〕側〔がわ〕 の カベ から 煙〔けむり〕 が 出〔で〕

て いる の を 従〔じゅう〕業〔ぎょう〕 員〔いん〕 が みつけ た 。 ＃

K0011290104010503 ＃ およそ 十〔じゅう〕 平〔へい〕方〔ほう〕 メートル の カベ を こがし た だけ で 消〔け〕しとめ た 。＃

K0011290101010504 ＃ 埋〔うめ〕立〔たて〕 地〔ち〕 に 男〔おとこ〕 の 変〔へん〕 死〔し〕体〔たい〕 ＃

K0011290104010505 ＃ 〔 船〔ふな〕橋〔ばし〕 〕 三〔みっ〕日〔か〕 午〔ご〕前〔ぜん〕 三〔さん〕 時〔じ〕 十〔じゅっ〕 分〔ぷん〕 ごろ ， 千〔ち〕葉〔ば〕 県〔けん〕 船〔ふな〕橋〔ばし〕 市〔し〕 日〔ひ〕の出〔で〕 町〔ちょう〕 一〔いち〕 〇〔れい〕 ， 第〔だい〕 一〔いち〕 コンクリート 会〔がい〕社〔しゃ〕 正〔せい〕門〔もん〕 前〔まえ〕 埋〔うめ〕立〔たて〕 地〔ち〕 の 臨〔りん〕海〔かい〕 道〔どう〕路〔ろ〕 ぎわ に 若〔わか〕い 男〔おとこ〕 が うつ伏〔ぶ〕せ に なっ て 死〔し〕ん で いる の を 通〔つう〕行〔こう〕 人〔にん〕 が みつけ 船〔ふな〕橋〔ばし〕 署〔しょ〕 に 届〔とど〕け出〔で〕 た 。 ＃

K0011290104010506 ＃ 新〔にい〕島〔じま〕 は 品〔しな〕川〔がわ〕 区〔く〕 大〔おお〕崎〔さき〕 郵〔ゆう〕便〔びん〕 局〔きょく〕 の 集〔しゅう〕配〔はい〕 人〔にん〕 と し て 勤〔きん〕務〔む〕 中〔ちゅう〕 ， 一〔いっ〕昨〔さく〕 年〔ねん〕 一〔いち〕 月〔がつ〕 十〔じゅう〕 二〔に〕 日〔にち〕 ， 品〔しな〕 川〔がわ〕

001 OPEN 3　OUTPUT-FILE

001 OPEN 4　SPECIAL-FILE

K001129 01 01010105 ｵｵﾀﾆ ｼﾞｭｳｺｳ ﾌｶﾞﾜ ｼｭｳｼｭｳﾖｳ ﾎﾞｻ

K001129 02 01040105 ﾐﾂﾙ ｺﾞｺﾞ ｲﾁ ｼﾞ ｼﾞｭｺｳ ｺﾞ ﾌﾝ ｺﾞﾛ ， ﾄｳｷｮｳ ｽｴﾀﾞﾍﾞ ｸ ﾔﾅﾊﾗ ﾁｮｳ ｲﾁ ﾚｲ ， ｵｵﾀﾆ ｼﾞｭｳｺｳ ﾌｶﾞﾜ ｺｳｼﾞｮｳ ﾉ ｷｶｲ ｺｳｼﾞｮｳ ﾖﾅﾐｶﾞﾜ ﾉ ｶﾍﾞ ｶﾗ ｹﾑﾘ ｶﾞ ﾃﾞ ﾃ ｲﾙ ﾉ ｦ ｼﾞｭｷｮｳﾖｳ ｲﾝ ｶﾞ ﾐﾂﾀﾞ ﾀ 。

K001129 03 01040105 ｵﾖﾋﾞ ｼﾞｭｺｳ ﾍｲﾎｳ ﾒｰﾄﾙ ﾉ ｶﾍﾞ ｦ ｺｶﾞｼﾀ ﾀﾞｹﾃﾞ ｹﾗﾄﾒﾀ 。

K001129 04 01010105 ｳﾒﾀﾃ ﾁ ﾆ ｵﾄｺ ﾉ ﾍﾝ ｼﾀｲ

K001129 05 01040105 ＋ ﾌﾅｶﾞｼ ＋ ﾐﾂﾙ ｺﾞｾﾝ ｻﾝ ｼﾞ ｼﾞｭｳ ﾌﾟﾝ ｺﾞﾛ ， ﾁﾊﾙ ｶﾝ ﾌﾅｶﾞｼ ｼ ﾋﾝﾃﾞ ﾁｮｳ ｲﾁ ﾚｲ ， ﾀﾞｲ ｲﾁ ｺﾝｸﾘｰﾄ ｶﾞｲﾍｷ ｾｲﾓﾝ ﾏｴ ｳﾒﾀﾃ ﾁ ﾉ ﾘﾝｶｲ ﾄﾞｳﾛ ｷﾞﾘ ﾆ ｳｶｲ ｵﾄｺ ｶﾞ ｳﾜﾌﾞｾ ﾆ ﾅﾘ ﾃ ｼﾝ ﾃﾞ ｲﾙ ﾉ ｦ ｻｲｺｳ ﾆﾝ ｶﾞ ﾐﾂﾀﾞ ﾌﾅｶﾞｼ ｼｮ ﾆ ﾄﾄﾒﾃﾞ ﾀ 。

K001129 06 01040105 ﾆｲｼﾞﾏ ﾊ ｼﾏｶﾞﾜ ｸ ｵｵｻｷ ｺｳﾋﾞﾝ ｷｮｸ ﾉ ｼｺｳﾊｲ ﾆﾝ ﾄ ｼﾃ ｷﾝﾑ ﾁｭｳ ， ｲﾘﾔｸ ﾈﾝ ﾋﾄ ﾂｷ ｼﾞｭｳ ﾆ ﾆﾁ ， ｼﾏｶﾞﾜ ｸ ﾌﾀﾀﾋﾞ ﾁｮｳ ｵ ﾊｲﾀﾘ ﾄｷｺｳ ﾃﾞ ｾﾝ ｴﾝ ｲﾘ ｸﾝﾚﾝ ｶｷﾄﾒ ｲﾘ ﾂｳ ｵ ｽｽﾝ ﾀﾞ ﾉ ｵ ﾊｼﾞﾒ ， ﾄﾞｳｽﾝ ｼﾞｭｳ ｲﾁ ｶﾞﾘ ﾏﾃﾞ ﾆ ｸﾝﾚﾝ ｶｷﾄﾒ ﾛｸ ｳｼ ｷｺｳｾﾝ ｺﾞﾋﾛｸ ｴﾝ ｵ ﾊｲﾀﾘ ﾉ ﾄｷｺｳ ﾃﾞ ｽｽﾝ ﾃﾞ ｷｿﾘ ﾚ ﾀ 。

K001129 07 01040105 ｲﾘﾔｸ ﾈﾝ ｼﾞｭｳ ﾆ ｶﾞﾂ ｼﾞｭｳﾊﾁ ﾆﾁ ， ｲﾘ ｾﾝ ﾉ ﾄｳｷｮｳ ﾁｻｲ ﾊ ﾋｺｸ ﾆ ﾁｮｳｴｷ ｲﾁ ﾈﾝ ﾛｸ ｶﾞﾂ ， ｼﾘｺｳ ｺｳﾖｳ ﾖ ﾈﾝ ｵ ｲｲﾜﾀﾘ ﾀ ﾉﾆ ﾀｲｼ ， ｸﾝｻﾂ ｶﾞﾜ ﾊ ｰ ｼﾘｺｳ ｺｳﾖｳ ﾊ ﾀｲ ｶﾞ ｶﾙ ｽｷﾞﾙ ｰ ﾄ ｼｽﾃﾁｮｳ ｼ ﾃ ｺｳｼ ， ｻｸﾈﾝ ｼ ｶﾞﾂ ｼﾞｭｳ ﾆ ﾆﾁ ， ﾆ ｾﾝ ﾉ ﾄｳｷｮｳ ｺｳｻｲ ﾊ ｸﾝｻﾂ ｶﾞﾜ ﾉ ｼｽﾃﾁｮｳ ｵ ﾐﾄﾒ ， ｲﾘ ｾﾝ ﾉ ｼﾘｺｳ ｺｳﾖｳ ﾉ ﾊﾝｹﾘ ｵ ﾊｷ ， ｱﾗﾀﾒ ﾃ ｻﾖｳｴｷ ｲﾁ ﾈﾝ ﾉ ｼｰｻｲ ｵ ｲｲﾜﾀｼ ﾀ ﾀﾒ ， ﾋｺｸ ﾊ ｺﾚ ｵ ﾌｸｸ ﾄ ｼﾃ ｼﾞｭｳｺｸ ｼ ﾃ ｲﾏｽ ﾀ 。

K001199 01 15081012 ｿﾉ ｶﾞ ﾀﾞｲｶﾞｸ ｵ ｿｿﾀﾞｮｳ ｻ ﾚﾙ ﾏﾃﾞ ， ﾂﾏﾘ ｺﾞ ｼｽﾃﾑ ﾆ ｲﾁﾊﾞﾝ ﾊｯｽﾙ ｼ ﾃ ｲﾀﾀﾞｸ ｼﾞｷ ｶﾞ ＋ ｶﾌﾄﾞｳ ｷ ＋ 。

K001199 02 15081012 ＋

K001199 03 15081012 ｽｴ ﾉ ｶﾞ ｹﾂｺﾝ ＋ ｼｽﾃﾑｼｮｸ ﾌｻﾙ ﾏﾃﾞ ｵ ＋ ｱﾝﾃｲ ｷ ＋ 。

K001199 04 15081012 ＋ ｿｼﾃ ﾌﾀﾀﾋﾞ ， ｺﾞ ﾌﾛｱ ﾀﾞｹﾞ ﾉ ｾｲｶﾂ ｶﾞ ﾊｼﾞﾏﾙ ＋ ﾛｸｺﾞ ｷ ＋ 。

K001199 05 15081012 ｺﾉ ｺﾞ ﾌﾛｱ ﾉ ｵ ﾗｲﾌﾙ ｻｲｸﾙ ﾄ ﾖｺﾞ ﾉ ﾃﾞｽ 。

K001199 06 15081012 ｿｺ ﾃﾞ ｵ ﾔｸ ﾆ ﾀﾘ ﾉ ｶﾞ ﾀﾞｲ ｲﾁ ﾉ ｾｲﾒｲ ﾎﾝ 。

K001199 07 15081012 ｶﾆｺｳ ﾉ ﾋﾞ ｶﾗ ， ｱﾅﾀ ﾉ ｾｲｶﾂ ｵ ﾎｼﾞｮｳ ｼ ﾃ ｸﾚ ﾏｽ 。

K001199 08 15081012 ﾀﾄｴﾊﾞ ， ﾔｽｲ ｶｹｷﾝ ﾃﾞ ， ﾏﾝｲﾁ ﾉ ﾊﾞｱｲ ﾊ ﾏﾝｷ ｳｹﾄﾘ ｷﾝ ﾉ 3 ﾊﾞｲ ﾎｼﾞｮｳ ＋ ｱｲ ﾉ ﾎﾝ ＋ ﾔ ， ﾏﾝｲﾁ ﾉ ﾎｼﾞｮｳ ﾊ ﾓﾄ ﾖﾘ ， ﾈﾝｽﾝ ﾌｴﾙ ﾊｲﾄｳ ﾃﾞ ｼｮｳｷ ﾉ ｶｹｷﾝ ｶﾞ ﾔｽｸ ﾅﾘ ， ﾁｮｷｸ ﾄ ｼ ﾃ ﾓ ｺｳﾘ ﾏ ＋ ﾄｸﾍﾞﾂ ｼｮｳｺｳ ＋

```
001 OPEN 1    INPUT-FILE
                                                                                        00001
K001129 01 01010105                                         ｵｵﾀﾆ ｼﾞﾕｳｺｳ ﾌｶｶﾞﾜ ｺｳｼﾞﾖｳ ﾎﾞﾔ
K001129 01 01010105                                    ｵｵﾀﾆ ｼﾞﾕｳｺｳ ﾌｶｶﾞﾜ ｺｳｼﾞﾖｳ ﾎﾞﾔ
K001129 01 01010105                             ｵｵﾀﾆ ｼﾞﾕｳｺｳ ﾌｶｶﾞﾜ ｺｳｼﾞﾖｳ ﾎﾞﾔ
K001129 01 01010105                     ｵｵﾀﾆ ｼﾞﾕｳｺｳ ﾌｶｶﾞﾜ ｺｳｼﾞﾖｳ ﾎﾞﾔ
K001129 01 01010105             ｵｵﾀﾆ ｼﾞﾕｳｺｳ ﾌｶｶﾞﾜ ｺｳｼﾞﾖｳ ﾎﾞﾔ
K001129 02 01040105                                         ﾐﾂｶ ｺﾞｺﾞ ｲﾁ ｼﾞ ｼﾞﾕｳ ｺﾞ ﾌﾝ ｺﾞﾛ , ﾄｳｷﾖｳ ｽﾐﾀﾞ ｸ ﾔﾅﾊﾗ ﾁﾖｳ ｲﾁ
K001129 02 01040105                                     ﾐﾂｶ ｺﾞｺﾞ ｲﾁ ｼﾞ ｼﾞﾕｳ ｺﾞ ﾌﾝ ｺﾞﾛ , ﾄｳｷﾖｳ ｽﾐﾀﾞ ｸ ﾔﾅﾊﾗ ﾁﾖｳ ｲﾁ ﾚｲ
K001129 02 01040105                                ﾐﾂｶ ｺﾞｺﾞ ｲﾁ ｼﾞ ｼﾞﾕｳ ｺﾞ ﾌﾝ ｺﾞﾛ , ﾄｳｷﾖｳ ｽﾐﾀﾞ ｸ ﾔﾅﾊﾗ ﾁﾖｳ ｲﾁ ﾚｲ , ｵｵ
K001129 02 01040105                             ﾐﾂｶ ｺﾞｺﾞ ｲﾁ ｼﾞ ｼﾞﾕｳ ｺﾞ ﾌﾝ ｺﾞﾛ , ﾄｳｷﾖｳ ｽﾐﾀﾞ ｸ ﾔﾅﾊﾗ ﾁﾖｳ ｲﾁ ﾚｲ , ｵｵﾀﾆ
K001129 02 01040105                          ﾐﾂｶ ｺﾞｺﾞ ｲﾁ ｼﾞ ｼﾞﾕｳ ｺﾞ ﾌﾝ ｺﾞﾛ , ﾄｳｷﾖｳ ｽﾐﾀﾞ ｸ ﾔﾅﾊﾗ ﾁﾖｳ ｲﾁ ﾚｲ , ｵｵﾀﾆ ｼﾞﾕ
K001129 02 01040105                     ﾐﾂｶ ｺﾞｺﾞ ｲﾁ ｼﾞ ｼﾞﾕｳ ｺﾞ ﾌﾝ ｺﾞﾛ , ﾄｳｷﾖｳ ｽﾐﾀﾞ ｸ ﾔﾅﾊﾗ ﾁﾖｳ ｲﾁ ﾚｲ , ｵｵﾀﾆ ｼﾞﾕｳｺｳ ﾌ
K001129 02 01040105                  ﾐﾂｶ ｺﾞｺﾞ ｲﾁ ｼﾞ ｼﾞﾕｳ ｺﾞ ﾌﾝ ｺﾞﾛ , ﾄｳｷﾖｳ ｽﾐﾀﾞ ｸ ﾔﾅﾊﾗ ﾁﾖｳ ｲﾁ ﾚｲ , ｵｵﾀﾆ ｼﾞﾕｳｺｳ ﾌｶｶﾞ
K001129 02 01040105               ﾐﾂｶ ｺﾞｺﾞ ｲﾁ ｼﾞ ｼﾞﾕｳ ｺﾞ ﾌﾝ ｺﾞﾛ , ﾄｳｷﾖｳ ｽﾐﾀﾞ ｸ ﾔﾅﾊﾗ ﾁﾖｳ ｲﾁ ﾚｲ , ｵｵﾀﾆ ｼﾞﾕｳｺｳ ﾌｶｶﾞﾜ ｺ
K001129 02 01040105           ﾐﾂｶ ｺﾞｺﾞ ｲﾁ ｼﾞ ｼﾞﾕｳ ｺﾞ ﾌﾝ ｺﾞﾛ , ﾄｳｷﾖｳ ｽﾐﾀﾞ ｸ ﾔﾅﾊﾗ ﾁﾖｳ ｲﾁ ﾚｲ , ｵｵﾀﾆ ｼﾞﾕｳｺｳ ﾌｶｶﾞﾜ ｺｳｼﾞﾖ
K001129 02 01040105       ﾐﾂｶ ｺﾞｺﾞ ｲﾁ ｼﾞ ｼﾞﾕｳ ｺﾞ ﾌﾝ ｺﾞﾛ , ﾄｳｷﾖｳ ｽﾐﾀﾞ ｸ ﾔﾅﾊﾗ ﾁﾖｳ ｲﾁ ﾚｲ , ｵｵﾀﾆ ｼﾞﾕｳｺｳ ﾌｶｶﾞﾜ ｺｳｼﾞﾖｳ ﾉ
K001129 02 01040105 ｶ ｺﾞｺﾞ ｲﾁ ｼﾞ ｼﾞﾕｳ ｺﾞ ﾌﾝ ｺﾞﾛ , ﾄｳｷﾖｳ ｽﾐﾀﾞ ｸ ﾔﾅﾊﾗ ﾁﾖｳ ｲﾁ ﾚｲ , ｵｵﾀﾆ ｼﾞﾕｳｺｳ ﾌｶｶﾞﾜ ｺｳｼﾞﾖｳ ﾉ ｷｶｲ ｺｳ
K001129 02 01040105 ﾞ ｲﾁ ｼﾞ ｼﾞﾕｳ ｺﾞ ﾌﾝ ｺﾞﾛ , ﾄｳｷﾖｳ ｽﾐﾀﾞ ｸ ﾔﾅﾊﾗ ﾁﾖｳ ｲﾁ ﾚｲ , ｵｵﾀﾆ ｼﾞﾕｳｺｳ ﾌｶｶﾞﾜ ｺｳｼﾞﾖｳ ﾉ ｷｶｲ ｺｳｼﾞﾖｳ
K001129 02 01040105 ｲﾁ ｼﾞ ｼﾞﾕｳ ｺﾞ ﾌﾝ ｺﾞﾛ , ﾄｳｷﾖｳ ｽﾐﾀﾞ ｸ ﾔﾅﾊﾗ ﾁﾖｳ ｲﾁ ﾚｲ , ｵｵﾀﾆ ｼﾞﾕｳｺｳ ﾌｶｶﾞﾜ ｺｳｼﾞﾖｳ ﾉ ｷｶｲ ｺｳｼﾞﾖｳ ﾐﾅ
K001129·02 01040105 ｼﾞﾕｳ ｺﾞ ﾌﾝ ｺﾞﾛ , ﾄｳｷﾖｳ ｽﾐﾀﾞ ｸ ﾔﾅﾊﾗ ﾁﾖｳ ｲﾁ ﾚｲ , ｵｵﾀﾆ ｼﾞﾕｳｺｳ ﾌｶｶﾞﾜ ｺｳｼﾞﾖｳ ﾉ ｷｶｲ ｺｳｼﾞﾖｳ ﾐﾅﾐｶﾞﾜ
K001129 02 01040105 ｳ ｺﾞ ﾌﾝ ｺﾞﾛ , ﾄｳｷﾖｳ ｽﾐﾀﾞ ｸ ﾔﾅﾊﾗ ﾁﾖｳ ｲﾁ ﾚｲ , ｵｵﾀﾆ ｼﾞﾕｳｺｳ ﾌｶｶﾞﾜ ｺｳｼﾞﾖｳ ﾉ ｷｶｲ ｺｳｼﾞﾖｳ ﾐﾅﾐｶﾞﾜ ﾉ ｶﾍ
K001129 02 01040105 ﾞ ﾌﾝ ｺﾞﾛ , ﾄｳｷﾖｳ ｽﾐﾀﾞ ｸ ﾔﾅﾊﾗ ﾁﾖｳ ｲﾁ ﾚｲ , ｵｵﾀﾆ ｼﾞﾕｳｺｳ ﾌｶｶﾞﾜ ｺｳｼﾞﾖｳ ﾉ ｷｶｲ ｺｳｼﾞﾖｳ ﾐﾅﾐｶﾞﾜ ﾉ ｶﾍﾞ ｶ
K001129 02 01040105 ﾝ ｺﾞﾛ , ﾄｳｷﾖｳ ｽﾐﾀﾞ ｸ ﾔﾅﾊﾗ ﾁﾖｳ ｲﾁ ﾚｲ , ｵｵﾀﾆ ｼﾞﾕｳｺｳ ﾌｶｶﾞﾜ ｺｳｼﾞﾖｳ ﾉ ｷｶｲ ｺｳｼﾞﾖｳ ﾐﾅﾐｶﾞﾜ ﾉ ｶﾍﾞ ｶﾗ ｹ
K001129 02 01040105 ﾛ , ﾄｳｷﾖｳ ｽﾐﾀﾞ ｸ ﾔﾅﾊﾗ ﾁﾖｳ ｲﾁ ﾚｲ , ｵｵﾀﾆ ｼﾞﾕｳｺｳ ﾌｶｶﾞﾜ ｺｳｼﾞﾖｳ ﾉ ｷｶｲ ｺｳｼﾞﾖｳ ﾐﾅﾐｶﾞﾜ ﾉ ｶﾍﾞ ｶﾗ ｹﾑﾘ ｶ
K001129 02 01040105 ﾄｳｷﾖｳ ｽﾐﾀﾞ ｸ ﾔﾅﾊﾗ ﾁﾖｳ ｲﾁ ﾚｲ , ｵｵﾀﾆ ｼﾞﾕｳｺｳ ﾌｶｶﾞﾜ ｺｳｼﾞﾖｳ ﾉ ｷｶｲ ｺｳｼﾞﾖｳ ﾐﾅﾐｶﾞﾜ ﾉ ｶﾍﾞ ｶﾗ ｹﾑﾘ ｶﾞ ﾃﾞ
K001129 02 01040105 ｽﾐﾀﾞ ｸ ﾔﾅﾊﾗ ﾁﾖｳ ｲﾁ ﾚｲ , ｵｵﾀﾆ ｼﾞﾕｳｺｳ ﾌｶｶﾞﾜ ｺｳｼﾞﾖｳ ﾉ ｷｶｲ ｺｳｼﾞﾖｳ ﾐﾅﾐｶﾞﾜ ﾉ ｶﾍﾞ ｶﾗ ｹﾑﾘ ｶﾞ ﾃﾞ ﾃ ｲﾙ ﾉ
```

④

K005887 02 10040111 ウマトﾞシ ニ チナミ ー カンエイ サン ハﾞシﾞユツ ー ト ウマ カﾞ アタﾞウチ スル ー マハﾞシ ケンコウ ー コウエン + +

K017917 04 05040109 ズﾞ トモ チユウオウ カラ チホウ ニ クラカﾞエ シ タ オオク ノ ウマ ハ ウマエ カﾞ カワル 。

K005887 02 10040111 ウマトﾞシ ニ チナミ ー カンエイ サン ハﾞシﾞユツ ー ト ウマ カﾞ アタﾞウチ スル ー マハﾞ

K001490 11 02040106 ウ フツリヨウ ノ チユウオウ アフリカ テﾞ クﾞンシﾞ セイケン カﾞ ウマレ タ チヨクコﾞ テﾞ アリ 、 サクネン ロク カﾞツ シﾞユウ ク ニチ 、 アルシﾞエリア テﾞ

K011134 06 06040311 シ カﾞ イウ トコロ ノ タﾞイ ニ シﾞ アハﾞンキﾞヤルトﾞ カﾞ ウマレ タ ノテﾞ アル 。

K008183 09 01040101 トウヨヨウ ト ウマレ 、 コﾞシﾞユツ サイ 。

K011664 07 14081012 ワカワカシイ カﾞクフウ カﾞ ヤキﾞ ニ ヨツ テ トウホク ノ チ ニ ウマレ ヨウ ト シ テ イタ 。

K004987 07 01040105 ヨ ニン オ シメコロシ タ アイチ ケン チタ クﾞン タケトヨ マチ ウマレ 、 モト コウイン: 、 サカキハﾞラ タケイチ + シ シチ + ハ シﾞユウ イチ ニチ ア

K011194 09 05040105 テ オリ 、 ハタケ ヤ サンリン ノ キユウリヨウ チタイ カﾞ ニ ウマレカワル ヒ モ チカイ 。

K017006 08 04030105 ウミ ノ オトコ タチ ノ ソレ カﾞ セメテ モ ノ イタワリ ト テﾞモ イウ ヨウニ 。

K014910 08 01040706 ト ハ イエ 、 ヘﾞイコク ナイ ニ アタラシイ キモ カン オ ウミ タﾞ シ 、 リンセン タイセイ ヘ ノ シﾞユンヒﾞ オ ウツカﾞシ テ イタ 。

K017898 05 05040111 ク チヨツト シ タ ハツケﾝ カﾞ ヒﾞミヨウ ナ エイキヨウ オ ウム コト モ アル ノ ハ 、 ムラロ トウセﾝ タﾞロ ウ 。

K004987 11 01040105 " カラ ナコﾞヤ チサイ ケイシﾞ ロク フﾞ テﾞ ヒラカ レ 、 ウメサﾞウ サイハﾝ カン ハ キンコ イチ ネン ロク カﾞツ 、' シツコウ ユウヨ コﾞ ネン + キユ

K001129 04 01010105 ウメタテ チ ニ オトコ ノ ヘン シタイ

K001129 05 01040105 レイ 、 タﾞイ イチ コンクリ ー ト カﾞイシヤ セイモン マエ ウメタテ チ ノ リンカイ トﾞウロ キﾞウ ニ ワカイ オトコ カﾞ ウツフﾞセ ニ ナツ テ シン テﾞ イル ノ

K004322 02 01040101 ル タメ 、 コレ タﾞケ テﾞ スイトﾞウ カイケイ ノ アナ オ ウメル コト ハ フ カノウ タﾞ カﾞ 、 シﾞミン 、 コウメイ リヨウトウ ハ '、 サﾞイケﾝ ノ

K003554 04 15081012 マエ + ヨコハマ エキ ニシクﾞチ + ヨコハマ タカシマ ヤ ウラ + カナサﾞワ + シﾞユツケン マチ カト + シスﾞオカ ニツコウ カイカンﾞ ニシトﾞナリ

K004822 05 05040307 カツコ ト シ タ アンテイ セイケン オ モツ ト イウ シﾞシン ノ ウラウチ カﾞ アル 。

K002507 10 10070103 "ナカﾞタ 5 2・ + サセホ 7 0 + + ウラカﾞ 3 8 + 1 ハコタﾞテ 4 6 + 1 + カウシケﾞ 6 0

K010307 18 10060111 ササメエキ ー + タﾞイ イチ フﾞ + タﾞン レイコ ウラシマ チカコ カミヤマ シケﾞル 1 1 + 3 0 + + 4 5 スホﾟ

K003572 02 10060111 スタンウイツク リチヤ ー トﾞ + ロンクﾞ + ニシノ サエ ウラノ ミツ ホカ

K006932 01 10060111 イツカ ー リチヤ ー トﾞ + ロンクﾞ + ニシノ サエ ウラノ ミツ

K002540 06 11040101 コトﾞク ナ タﾞンシﾞヨ フタリ カﾞ 、 コノ タﾞイ トシ ノ ウラマチ ノ サカハﾞ テﾞ クﾞウセﾝ ニ シリアイ 、 ソシテ ムスハﾞ レ テ ユク 。

K003554 03 15081012 ワ + エキ キタクﾞチ キﾞンサﾞ テﾞハﾟ ー ト トナリ + ウラワ + チハﾞ ケイセイ カイカン ウチ +

K005857 07 05040303 セﾝキ ト ホホﾞ トﾞウカﾞク ノ セン ヨンヒヤク オク エン ノ ウリ アケﾞ ハ カクシﾞツ テﾞ 、 ココ トウフﾞン ノ カン ハ イチ ワリ ロツ フタン ハイトウ オ

K011724 03 03040607 、 イマ ノ サトウ シコシヨウ ノ ー マワリ ー ニ ウロウロ スル ヒト ハ イ テ モ ー ソハﾞ チカク サカエル ヒト ー ハ イ ナイ 、 ト イウ

K012962 02 01040101 コクミン ケイザイ カ アンテイ テキ ナ セイチョウ オ ツヅケル タメ ニ ハ 、 ユウコウ シヨウ ト キヨウエウ ノウリョク ト

K008658 05 01040702 〃 、 ソレン ノ イト ハ コノ ユウコウ カンケイ オ イチジ テキ ナ ー ム ー ト ー ニ オワラ セル コト ナク 、 ケイザイ 、 ボウエキ 、

K004357 11 01040206 ハ イツポウ テキ ニ ボウエキ コウシヨウ ノ シヨウキヨウ オ コウヒヨウ シ タ ガ 、 コレ ニ ツイ テ ハ

K006845 06 01040111 カシヤマ サント 、 トテモ ニ テイタノ カ インシヨウ テキ ダツタ 。

K010112 03 01040706 ツケタ ホクハク ノ コウカ ニ ツイテ 、 カナリ カイキ テキ ニ ナツ テ イル コト オ シメス モノト ウケトラ レ テ イル ガ 、 シシツ 、 セン

K005857 05 05040303 サクネン 、 カカク テキ カンリ ホウ オ シツシ 、 デミング シヨウ オ ウケ タ カ 、 コトシ ハ トヨタ リ

K013126 05 04040105 カンシユウ テキ ナ ケンリ デ 、 ミンポウ テモ ー イリテイ ケン ハ エク チホウ ノ カンシユウ ニ シタ

K014511 06 10040303 テ モ 、 シヨウヒ シヤ ブツカ ノ シヨウシヨウ ハ キホン テキ ニ ハ チンアゲ + キキヨウ シユウエキ ノ アツハク ト ナツ テ 、 カブタカ ノ テキ

K005840 08 04040106 ワレワレ ハ セントウ オ キヨクテ テキ ニ カキル コト オ ノゾン タ カ 、 セツキヨク カ スル シンリヤク オ ケッタイ シ

K002540.05 11040101 ソレ モ キンタイ テキ ナ コウソウ ヒル ノ タチナラブ ニユ ー ヨ ー ク ノ マンハツタン タ 。

K013126 06 04040105 シユウヨウ ナ ケンリ ト サ レ テ イル カ 、 キンタイ テキ ナ シヨユウ ケン ト ハ マツチ シ ナイ タメ 、 セイリ サ レル ケイコウ ニ アル 。

K011526 04 01040101 カ シユウテン テ 、 ブツカ タイサク ハ タイニ キ テキ ニ ナツ テ イル 。

K010112 01 01040706 シ タ ト ハ タンシテ オラ ズ 、 ソノ トキ ノ クタイ テキ ソチ ニ ツイ テ モ ケッテイ シ テ ハ イ ナイ カ 、 ケンサイ デ ハ ヘイワ ヘ ノ シ

K008658 05 01040702 シ 、 カイコウ ノ フンヤ デモ デモ タヤ クタイ テキ ナ カンケイ ト シ テ ヒトツ ヒトツ シツ オ ムスハ セ テ ユキ タイ ト イウ テン ニ アル

K016036 03 01040105 カツコウ カウ ニ モ クタイ テキ ナ カイケツ サク ハ ナク 、 シヨクン ノ リヨウシキ ニ マツ ホカ ハ ナイ ー ト ノベ タ

K016945 03 15081012 コンゲツ カラ ケイカク テキ ニ チヨチク シ テ シブヘン デ ネンキン オ サクロ ウ ナンテ イイ タス ン タ モノ ウフフ

K007834 02 01040105 カ ホイラ ー オ タイ テ イル ヨリ モ スベツト ケイザイ テキ デ 、 トウシヨウ ノ シサン デ ハ スツリヨウ ヤ ロウム ヒ 。 トウリヨク ヒ 、

K010110 06 01040706 ホクハク テイシ ハ コレ ニ ヨツ テ カイコウ カ ヨリ コウカ テキ ナ ヤクワリ オ ハタシ 、 ヘイワ ヘ ノ ミトオシ カ カイセン サ レル ト スル キヨウサン

K001987 07 06040407 ヨウスルニ 、 シコウト ノ セカイ ト シ テキ ナ セカイ ノ コントウ カ ミ ラレル ノ デ アル 。

K001987 09 06040407 セカイ オ タノシム タメ ノ タイ キヨウシ ト ハ シヤカイ テキ シツ カ コトナツ テ クル ノ デ アル 。

K001987 10 06040407 、 ニホン シン ノ オ シヨウガツ ニ ナイ ヨウナ シヤカイ テキ シユウヨウ セイ カ アル ノ デ アル 。

K008212 07 01040206 イチ 、 ソレン ト モンコル ハ シヤカイ シユキ テキ コクサイ チシノ ケンソク ニ モトツイ テ セン メン テキ ナ キヨウリヨク カンケイ ト ユウコウ テ

K004357 07 01040206 ク + キユ ー ハ ボウエキ ニ ツイ テ ケンサイ シヨホ テキ ニ マトメ ラレ タ ハナシアイ ニ ヨル ト 、 ボウエキ カク ハ サクネン ニ クラベ イクラカ

K011526 16 01040101 ウ リヨウキン カ トウシ ニ アカラ ナイ ヨウ 、 シキ テキ ナ サ オ ツケル ヨウニ シ テ イル 。

K009796 03 01040105 テ + キソ シシツ ハ カメタ ヒコク トウヨウ シツシツ テキ ニ ハ イホウ セイ オ ソキヤク スル モノ 、 アルイハ コウム シツコウ ボウカイ サイ カ

⑥

K026041 06 05040107 ハ ， イマ ノ ニホン カイ イツタイ カ オオキナ サンカ゛ク オ モツ タイリク タ゛ツ タ 。
K075113 06 15081012 － アントロマック － オ シノク゛ シ゛シン オ モツ テ ニツセイ ケ゛キシ゛ヨウ カ カントウ キヨヘン
K088656 01 10040111 ウ オ メツシ テ イル アメリカ ハ ， タイヘン ナ シ゛シン オ モツ テ イ テ ， ステ゛ニ ソレン オ オイコシ タ ト イウ ホコリ オ モツ テ イル －
K005286 01 01040101 ル ヨウタ゛ カ ナカ゛イ メ テ ミ テ シテ オク ノ シ゛ンコウ オ モツ チユウコク゛ ユウコウ カンケイ ハ フカメル ヘ゛キ テ゛ アル 。
K067280 11 01040111 ウ ヤブ゛レ トE イウ ヘ゛キ ， ヒラカ レ タ カンシ゛ユ セイ オ モツ トコロ ニ ， コノ サツカ ノ フシキ゛ ナ ミリヨク カ アル 。
K024575 10 10040303 キヨウト゛ウ シヨウセン ハ モツト コウ テキ ナ セイカク オ モツ タ キカン タ゛ ト スEウ 。
K004822 05 05040307 ヘイ ノ ハ゛ランス ノ ソE ニ カツコ ト シ タ アンテイ セイケン オ モツ ト イウ シ゛シン ノ ウラウチ カ アル 。
K001987 10 06040407 イ セイカツ ニ スイ テ ， ハツキリ コトナル フタツ ノ セカイ オ モツ テ イル ヒトビト ニ トツ テ ハ ， ニホン シ゛ン ノ ス シヨウカ゛ツ ニ ナイ ヨウナ シ
K085494 01 14081012 ト゛クタ゛ン センコウ ， セキニン オ モツ テ コノ ナンキヨク ニ アタル キヨウリヨク ナ ノ ホカ ハ ナイ +
K078759 08 14081012 ハ + トクニ サイコ゛ウ ヨシノスケ ハ トツ ノ ヨウ ナ タイサク オ モツ テ イル ノ テ゛ アロ ウ カ +
K081783 01 01040105 シヨ ウシテ゛カ チユウサ゛イ シ゛ヨ ヘ ワカイ オトコ カ タントウ オ モツ テ オシイリ ， オウタイ ニ テ゛ タ ヤマEト ケンタロウ シ゛ユンサ + コ゛ ハチ +
K064536 13 11040111 ヒ゛ コウホウ フ゛ テ゛ ハ － ワカイ ラクコ゛ カ カ゛ チカラ オ モツ テ キ タ コト 。
K013498 01 15081012 + ツマミ オ モツ テ サユウ ニ ウコ゛カシ テ クタ゛サイ
K056447 08 01040105 ケ゛ンサ゛イ タイトウ ク ウエノ ニ ニ ケン ノ ハンハ゛イ テン オ モツ テ イル 。
K064950 04 08030105 ツ タ コレ ラ ノ イツ セイ タテ ハ ， イマ トツ゛ンナ ナヤミ オ モツ テ イル ノ タ゛ロ ウ カ 。
K002540 09 11040101 シラヒ゛ヨウシ テ゛ アル シスヘ゛カ コ゛セ゛ン ハ セイカツ ノ ニオイ オ モツ シ゛ンフ゛ツ テ゛ス 。
K081208 08 15081012 + ハツヒ゜ヨウ + トウセン シテ ハ シヨウヒン ノ ハツソウ オ モツ テ ハツヒ゜ヨウ ニ カエ サセ テ イタタ゛キ マス 。
K041367 05 15081012 ハツヒ゜ヨウ ケ゛ンセイ チユウセン シヨウヒン ハツソウ オ モツ テ ハツヒ゜ヨウ ニ カエ マス
K046112 07 04040305 ニ ハ チカ テツ ハ アル カ － － ケ゛ンシ ハ゛クタ゛ン オ モツ タ チユウキヨウ ハ ニホン オ オソウ コト ハ ナイ カ － － コウキ゛ヨウ スイシ゛ユン
K053955 04 01030105 チ ヨ ， シンルイ ノ イエ ニ チフスキン ノ ツイ タ ハ゛ナナ オ モツ テ イキ ， チヨウチフス オ シユウタ゛ン ハツセイ サ セ タ コト オ シ゛キヨウ シ タ
K017917 09 01040105 スオサワ ヒロノリ サン + ニ コ゛ + ホウ ニ ホウチヨウ オ モツ テ オシイリ ケ゛シュク ニン ノ チユウカク゛イカ゛ク ホウ カ゛クフ゛ サン ネン ， ミウラ
K088656 01 10040111 テ イ テ ， ステ゛ニ ソレン オ スイコシ タ ト イウ ホコリ オ モツ テ イル － ト イウ 。
K070627 08 06040107 － ユウコウ － ノ オ ミヤケ オ モツ テ コ゛ カツ トウカ ニ カイロ ニ ツイ タ コスイキ゛ン シユシヨウ オ シケ゛キ シ ハ シ
K046156 01 11040407 ソレ ハ コンニチ ノ サツカ ハ ， セツクス シカ カタル モノ オ モツ テ イ ライ ， ト イウ ヨウナ カンカ゛エ カ ハイコ゛ ニ アル ニ チカ゛イ ナイ 。
K012961 09 01040101 キ トシ テ゛ アル ト カタク ケツイ シ ， コンシン ノ ユウキ オ モツ テ コノ モンタ゛イ ニ タイシヨ シ テ マイル カクコ゛ テ゛ アル 。
K041948 03 01040107 シコ゛ト ニ ユウキ オ モツ テ トリクム ニ ハ ， シンヘ゜ン ハ セイケツ テ゛ ナケレ ハ゛ ナラ ナイ 。