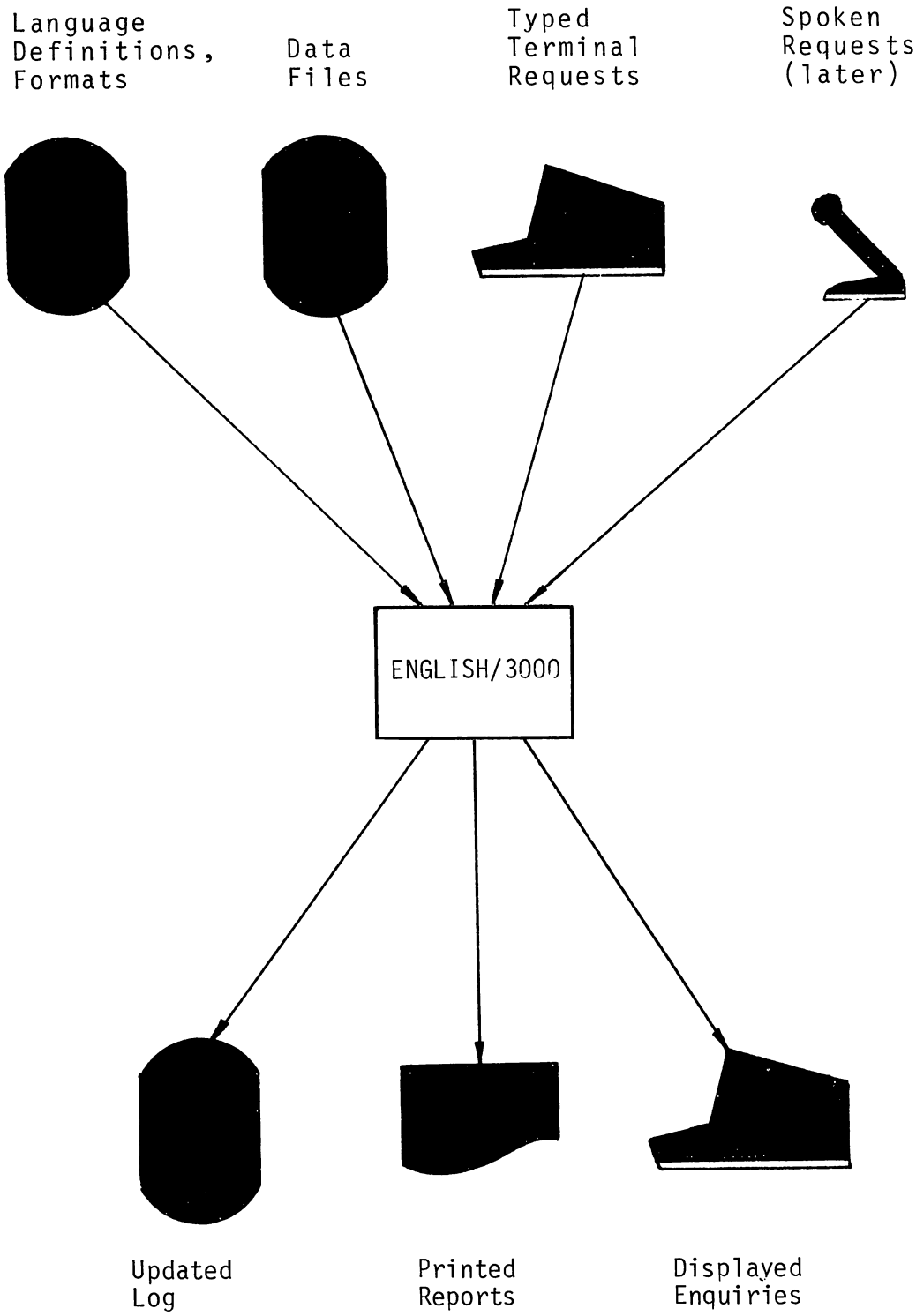
E N G L I S H / 3 0 0 0

A NATURAL LANGUAGE ON A MINI-COMPUTER

Doug Peckover
D.P. Concepts Inc.
98 Perodeau
Vaudreuil
Québec J7V 5V5
(514) 455-1373

Tuesday G-2 - 01

ENGLISH/3000 SCHEMATIC



WHAT & WHY?

"Hardware vendors could go a long way toward eliminating the problem of computer-caused unemployment if they were somehow able to achieve a technological breakthrough in natural computer languages. Development of a powerful, easy-to-use natural language would make computing systems available for the first time to a large class of unskilled users who would otherwise find the systems forever intimidating and inaccessible".

Alvin Toffler - author of "Future Shock"

What Mr. Toffler said is equally applicable to the thousands of "unskilled"managers who already have access to computers, but find them "intimidating and inaccessible".

There is presently a furious race to produce the world's first Voice Recognition Unit (VRU) that would understand our most natural language - our native tongue. Everyone from the heavies at the IBM T.J. Watson and Yorktown Heights labs as well as the Bell labs to fairly small micro and terminal manufacturers are producing encouraging, but crude VRU's. The general feeling is that a fully "unrestricted" VRU is still 8-20 years away. However, there are many enormous benefits to natural languages that are available right now. ENGLISH/3000 will shortly permit the HP 3000 users to take the first of three phases to a fully unrestricted natural language:

Phase 1 Accept a TYPED terminal command in a NATURAL (unrestricted) format. Interpret and execute the command. ("PLEASE SHOW ME ALL FINAL PRODUCTS THAT USE A 56421 BOLT" - this can be abbreviated).

Phase 2 Accept a SPOKEN voice command in an ARTIFICIAL (restricted) format. Interpret and execute the command. ("SHOW FINAL PARTS 56421").

Phase 3 Accept a SPOKEN voice command in a NATURAL format. Interpret and execute the command. ("PLEASE SHOW ME ALL FINAL PRODUCTS THAT USE A 56421 BOLT").

Phase 3 is probably 8-10 years away. Phase 2 is 3-5 years away. Phase 1 is available now.

There are many advantages to simulating VRU's. These include:

1. User training for enquiries and reports would be reduced to a minimum. Instead of having to study your documentation and running your programs, they would explain what they require just as they would to another person.

2. Slight variations in how a request is worded could format the information in many different ways. For example, "Display the parts list..." and "Display the costed parts list..." could indicate different requirements for two separate departments.

3. You will be able to start designing for the second and third stages that will have the most profound impact on your organisation. What we learn will not only prepare you now for VRU's but will also simplify the selection criteria for them as they come onto the market.

When the information requested is ready to print, we need a powerful report generator to produce terminal enquiries and printed reports. The benefits provided by ENGLISH/3000 include:

1. Traditional programming costs for most enquiries and reports would be slashed by up to 90%. The net increased productivity to your programming staff could double their output.

2. Your managers and users will be able to write their own enquiry and report formats. This would remove much of the load traditionally placed on the data processing department.

As with all major projects, specific design goals were set for ENGLISH/3000. The following pages outline some of the prime considerations that specified the framework on which the product was developed.

EASE OF USE

There can be no easier way to train a manager or user how to use a new system than by permitting him to use his natural language, whether it be English, French, Spanish, and so on.

1. No matter how complex the data structures are or how involved the programs get, when it comes down to a one on one enquiry, the user can communicate with the computer on an equal basis. There are no programs to memorize, no keys to remember, no rules to look up.

2. The training for enquiries and reports becomes the responsibility of the System Manager as he is the one to determine who needs what. The reduced number of conventional enquiries and report programs will make new systems easier to teach. In addition, ENGLISH/3000 could significantly reduce the sales effort needed by software OEMS by giving the System Manager these additional customizing facilities. This ease of use could even lead to a lower support cost for many installations.

3. The user can mix dependent demands, such as bill of materials with independent demands, such as customer orders, without the need to understand their relationship or dependancy on the data base(s).

4. Special considerations have been taken in designing the report writer. The assumed technical knowledge of the System Manager is EDIT/3000 and QUERY/3000. An ENGLISH/3000 format (or procedure) compiles itself automatically the first time it is used after a modification has been made. When you receive each new release of ENGLISH/3000, any modifications required to your source code (if any) will be automatically made and the format will be recompiled.

5. The formats that drive each request in ENGLISH/3000 are coded in a non-procedural way. This further reduces the need for amount of technical training required to code formats.

6. ENGLISH/3000 allows full comments - both at the heading and line level. The product will soon have an automatic flowcharting facility. This will document any format, showing loops, labels, comments, and so on.

SPEED CONSIDERATIONS

Most of the advantages of a natural language would be thrown out the window if the computer took too much time to process the request.

1. As a design goal, the average response time required to accept, interpret and begin executing the command (start displaying the enquiry or start STREAMING the report) will be 2 seconds. A tuning aid is provided that will help you improve the efficiency of your enquiry and reporting load on the computer.

2. ENGLISH/3000 supports KSAM indices for IMAGE master data sets. This eliminates the need for a serial read through the entire data set and the subsequent sort. A utility for maintaining these indices is included with ENGLISH/3000.

3. All reports are automatically STREAMed from the terminal. This not only frees the terminal for the next command but also permits the system to lower the report priority and queue the reports if necessary. At the user's option, the report can be flagged to only start after hours.

4. The source code used to describe the format is automatically compiled and stored when it is first used. All subsequent requests to the same format result in a much faster execution time.

5. Most enquiries and reports on your computer (say - 50% of your workload) would be handled by one single very efficient program that can be locked in memory. This would greatly reduce the system swapping by using the re-entrant facilities in the 3000. System thruput would be increased, improving response time, not only with ENGLISH/3000 but also your other application programs. This may reduce the need for hardware upgrades and possibly reduce the need for after-hour operations.

SECURITY CONSIDERATIONS

If natural languages will enable anyone to communicate freely with the computer, then a necessary part of our design objectives must be a complete review of conventional security methods.

1. ENGLISH/3000 comes with security at two levels. The first is the security provided with IMAGE. The second has been taken from a system designed for a large military supplier that had to be very secure. The System Manager identifies everyone in the company in a tree structure. The ENGLISH/3000 formats are then assigned to the users on a need-to-know basis. ENGLISH/3000 permits the user or any of his superiors to have full access to the format. However, no one below the user in the structure has access to the format. Invalid attempts to run a format are handled conversationally by telling the user see the department head that "owns" the format ("Please see Bob Smith..."). The error is also logged for that department head's attention.

2. If the user makes a specified (default 4) number of consecutive errors in this manner the terminal is jammed and must be freed by the operator. This makes the system secure if dialup terminals are used.

3. The System Manager has additional facilities for reports. He may choose to receive a log of all occurrences of a certain report, who ran it, when, and for what reason. He may also have a special security label precede any report. This will contain CONFIDENTIAL in large letters as well as the user requesting the report, delivery instructions, and so on. He may also have a system-generated serial number label each page of the report and log it's occurance.

4. After a pre-defined period of time, ENGLISH/3000 will do a timeout and re-request the user identify himself. This means that a user who has forgotten to sign off will not leave his terminal in a ready state.

5. As we move closer to VRU's, the user's password will be needed to identify his speech patterns. Eventually, the speech patterns themselves will be used to identify the users.

EVENT LOGGING

There are many useful byproducts of ENGLISH/3000 that are available from the logging facility. These will most likely be expanded to meet the changing needs of the client base requirements.

1. One of today's trends is to treat computers as profit centers. Departments, users, clients, and budgets are charged a portion of the costs on a flat rate or as-used basis. To facilitate this, ENGLISH/3000 has a sophisticated logging facility that enables you to charge according to the enquiry and reports used. You may charge by any combination of the following: lines or pages that are displayed or printed, CPU seconds used, elapsed time, disk reads and/or copies requested.

2. Sensitive reports, such as price lists, can be logged with their serial number printed at the bottom of each page. This then enables you to keep a journal of who has what version of what report.

3. The log can be used to analyse who is using what facilities. This will enable you to use the tuning facility to improve the efficiency of ENGLISH/3000.

4. The log can note unsuccessful attempts to run a program ("Bob Smith tried to run the PRICE LIST and was told to see Bill Boss"). Arrangements can then be made to see if Bob should have access to this request.

5. ENGLISH/3000 will use the logging facility to note requests that it could not interpret. This may be periodically analysed so that changes can be made to the language definitions. Studies indicate that this improves the chances of ENGLISH/3000 understanding a request on the first attempt from 90% up to 98% of the time.

6. At a later stage, ENGLISH/3000 can be modified to log certain data from key reports (such as monthly totals). These can in turn be reported on by selectively analysing the log file. These enhancements will be defined by user response to surveys.

DATA STRUCTURES

The usefulness of a natural language facility should not be limited to crude reports that must be used to output the information.

1. ENGLISH/3000 will support any combination of IMAGE data bases (multiple data sets), KSAM and MPE files as well as remote and local computers.

2. While not specifically aimed at the manufacturing users, there will be several data structures available that will be of particular use to manufacturers. Bill of material and where-used recursive structures to a specified number of levels will be supported. In addition, the format can specify that the intermediate sub-assemblies should not print, and only the final level should print. This is commonly required to show the raw materials for a given product or show which final assemblies use a particular sub-assembly or raw material.

3. Generic names and keys will be available for record selection, conditional printing and conditional branching.

5. Optional data set and file locking will enable the enquiry or report to be completed without any unexpected events.

PRODUCT SUPPORT

The intent of ENGLISH/3000 is to not only provide you with a powerful facility now, but to eventually accept requests via VRU's in a fully unrestricted manner.

1. Normal support is provided to continuously upgrade and improve ENGLISH/3000 for VRU's as well as to fix bugs and design oversights. By purchasing ENGLISH/3000, you are guaranteed an upward compatible path that will lead you and your organisation to the world of fully unrestricted natural languages with VRU's. The newsletter (described below) will help guide you on the selection of VRU's as they become available.

2. Extended support will shortly be introduced. This will provide you with a phone-in consulting service as well as a revolutionary new service that we believe is new to the industry. For installations that do not have programmers or installations that experience peak loads, we will offer a phone-in programming service. Your enquiry or report will be designed, coded, tested, and documented on your computer within 24 hours. (The request must be within the design limitations of ENGLISH/3000 and we must have access to your computer via a good dialup line). The request will be billed on a time and material basis with an agreed to ceiling.

3. D.P. Concepts Inc. will publish a monthly newsletter to keep you, the user, up-to-date and informed. Featured will be user tips and recommendations, progress reports on the VRU developments from the major labs and hardware vendors with product testing and recommendations, developments and standards from the American Association for Artificial Intelligence, notes on evolving natural language applications, and user survey forms for proposed enhancements. This newsletter will be sent to all ENGLISH/3000 users with either normal or extended support.

SUMMARY

Besides the more obvious reasons for buying a natural language facility (with a powerful report generator) three more important reasons exist:

1. A well designed natural language facility can actually pay for itself and save you money in the following ways:

   - reduce training times
   - reduce programming times
   - reduce documentation times
   - reduce lead times for new programs and modifications
   - charge users and departments for resources used

2. In the short term, natural languages will be used mainly for enquiries and reports. Eventually, the natural languages will be used to accept input as well. In addition,development of artificial intelligence on computers will permit the user to try more powerful requests, while requiring a reducing knowledge of how the computer works. ENGLISH/3000 will pursue these trends.

3. A well designed natural language should be able to accept input, interpret, and produce enquiries and reports for virtually any language. The only limitation appears to be whether the language has a terminal and printer that supports the character set for the particular language. Within this limitation ENGLISH/3000 can be converted to any other language.

PLEASE SHOW THE CURRENT LEVELS FOR A 1736-90.

... could execute the following format ...

```
REPORT
<<    Enquiry Name:    INVENTORY LEVELS
<<    Author       :   Manual Labour
<<    Last Changed:    23 Jan 81

<< This format displays the current inventory levels for
<< the selected part number by branch. The total for all
<< branches is shown at the end of the enquiry.

<< Steps:    - Read PART-NO from ITEM-MASTER
<<           - Use PART-NO to set up STOCK-LEVEL details
<<           - Read and print each detail
<<           - Use BRANCH to get DESC in BRANCH-MASTER
<<           - Print "*" if BRANCH-MASTER STATUS is "S"

IM MEANS ITEM-MASTER                    << Abbreviate "ITEM-MASTER"
SL MEANS STOCK-LEVEL                    << Abbreviate "STOCK-LEVEL"
BM MEANS BRANCH-MASTER                  << Abbreviate "BRANCH-MASTER"

H1,"Inventory Status Enquiry",26        << Enquiry Heading
H1,DATE-TIME,66                         << Set up Date-time stamp
H2,"Part Number",22,SPACE 2B            << Column Headings for PART-NO
H2,"Branch",37                          << Column Heading for BRANCH-NO
H2,"On Hand",52                         << Column Heading for QUANTITY
H2,"Status",66,SPACE 2A                 << Column Heading for STATUS flag

D1,IM.PART-NO,20                        << Print PART-NO from ITEM-MASTER
D2,SL.PART-NO(IM.PART-NO),DUMMY         << Set chain read in STOCK-LEVEL
D2,BM.DESC(SL.BRANCH-NO),41             << Print DESC from BRANCH-MASTER
D2,SL.QUALTITY,52,"ZZZ9.99"             << Print QUANTITY from STOCK-LEVEL
D2,IF,"S",EQ,BM.STATUS,"*",63           << Print "*" if STATUS = "S"
TF,"Total:",37,SPACE 2B                 << Footing literal
TF,QUANTITY,52,"ZZZ9.99"                << Print total at end
END
```

... which would display the following results ...

Inventory Status Enquiry WED, MAR 18, 1891, 9:04 PM

Part Number	Branch	On Hand	Status
1736-90	NEW YORK	142.00	
	DALLAS	9.50	*
	CHICAGO	14.00	
	Total:	165.50	