NATURAL LANGUAGE COMPUTING: THE COMMERCIAL APPLICATIONS

Tim Johnson

Ovum Limited
44 Russell Square
London WC1B 4JP

This is an edited version of the executive summary of Ovum's 1985 report of the same name.

THE IMPORTANCE OF NATURAL LANGUAGE PROCESSING

A new oportunity

The computer industry is in the early stages of a revolution. Its capability for processing natural languages will advance dramatically over the next few years. As a result, computer systems which deal with human language will begin to play a central role in the industry. Computers will become significantly more accessible to many more people, and many tasks which are now wholly manual will be handled electronically.

Teaching computers to understand English or any other natural language has been a dream for many years. But it has always been unrealistic, except in limited circumstances. The sheer complexity of natural language, and the wealth of knowledge which people unconsciously draw on to use or understand it, has defeated many attempts at computerisation. Now the situation is beginning to change. Solutions are being found to many of the problems, and new demand is springing up for facilities which only natural language processing (NLP) can provide.

The exponential increase in the computer power and memory available at any given price level is obviously one of the forces which will drive the revolution. NLP is a heavy user of both, and some of its most important applications will depend on the availability of tens of megabytes of memory and tens of mips (millions of instructions per second) of processing power at prices which would have been unthinkable a few years ago. As far as using the power is concerned, the development of expert systems and other artificial intelligence techniques is providing the software tools which make NLP easier to implement.

But NLP technology is making important advances of its own as well. New grammars are being developed to provide more complete coverage of language and more efficient computation. Reliable computer recognition of spoken words is rapidly being achieved, and will probably make this the most popular channel of communication with computers in future. Researchers are learning to limit the domains in which systems operate, and identify the strategies which will make them successful.

The changing computer environment is also helping to create increased demand for NLP. Desktop micros and data communications make it feasible to offer computer power to many more people, if they have the skills to use it. Thanks to word-processing and electronic mail, the volume of text which is generated and used within an electronic environment is growing rapidly, and provides an obvious target for computer processing. The increasing size and complexity of databases and other software requires more flexible, "wide bandwidth" interfaces for communication with their users.

NLP is one way of meeting these new demands — and the evidence that the technology to do it is emerging has been growing even while this report was being prepared. Most important of all are the recent rapid advances towards realisation of a practical talkwriter — a machine to accept human speech and transcribe it into machine-readable form, word-processor text perhaps, or commands to a database. Voice recognition systems which can accept spoken commands in the form of isolated words or phrases are already well known. In 1984 IBM demonstrated a talkwriter which can transcribe complete business letters, with a vocabulary of 5,000 words. There are at least seven more projects working towards similar but more advanced systems in the USA, the UK and Japan, and the first sample product will probably be delivered in 1986.

Talkwriters will be the main gateways to the natural language networks of the future, and the recent rate of progress suggests they will be available sooner than anyone dared to predict a few years ago. The gap between demonstrating a prototype which performs well in the laboratory and achieving acceptable operational performance is still large. There will be some years of test and trial before the samples of 1986 are ready to be used in a wide range of routine applications. But when it arrives, the ability to provide flexible

voice input to the whole gamut of applications, from word-processing to databases, will give NLP a central role in the computer industry.

The accelerating progress of NLP is confirmed by advances in several other applications besides the talkwriter. As recently as 1983, content scanning — using a computer to identify the content of messages or other semi-formatted texts — was the subject of just a handful of pioneering academic projects. In mid-1985 Cognitive Systems, a company specialising in NLP, announced the first content scanning product. Called "Atrans" it reads international money-transfer telexes. Systems like Atrans could be saving significant sums of money for banks and other organisations with large volumes of message traffic in the near future.

In other cases, the rate of progress of NLP applications that have been operational for years is increasing in 1985. Artificial Intelligence Corporation began developing its "Intellect" natural language interface for IBM databases in 1976, and achieved its first commercial installation in 1981. From 1981 to 1984 it averaged about 70 installations a year; since mid-1984 the rate has doubled. At the same time, several competitors have entered the market, so the total number of NL interfaces has shown an even sharper increase — although the absolute numbers are still small by comparison with the potential.

There has been a similar change in the market for machine translation (MT) — the use of computers to translate between human languages. Some big mainframe-based systems have been in use since the late 1960s, doing crude translations of Russian technical texts for the US Air Force for example. Only in 1984 did a new generation of companies such as Logos, ALPS and Weidner seriously begin to offer systems for the office market; the pace of activity in MT is now accelerating.

NLP may even be called in to revive the flagging micro boom. Several companies are offering different varieties of NL interfaces for personal computers; an enthusiastic reviewer has already declared that one of them ("Clout" from Microrim) "is destined to become a software best-seller." Even if that accolade is premature, someone is bound to come up soon with the precise mix of power, user-friendliness and marketing ability needed to make a big impact.

The signals all point to the same conclusion; the capability of providing effective NLP for a wide range of applications is now in reach. If this conclusion is accepted, then the importance of NLP to the computer industry cannot be denied. Once people are used to using natural language effectively to access computer resources, they are unlikely to go back to older methods. When MT or content scanning systems can provide more accurate results at lower costs than purely manual methods they will rapidly take over the market. This report aims to identify the areas where NLP is expected to be successful, and to suggest how people and organisations can make use of the opportunities it will create, as users, suppliers or investors.

The task of natural language processing

The objective of natural language processing is to take inputs in human language and extract from them something which is meaningful to a computer. The inputs can be typed "strings" of words, or spoken "utterances," entered on a keyboard, read into the system from a word-processor file, or spoken into a microphone. The meaning which is extracted may be commands to access a database; a grammatical structure which can be used to produce a translation; or, in the case of a talkwriter, the ASCII characters equivalent to the user's spoken words. Providing the means for computers to generate replies to queries and other outputs in natural language form is another branch of the same technology.

Almost any computer system can be set up to respond to single word commands, entered by a keyboard or spoken via a voice recognition system. The need for natural language processing begins when the words in the input string or utterance combine to produce meaning. At this point there are two alternative strategies available to the designer of an NLP system — emphasising either syntax or semantics. The syntax of a language is a description of its structure; the way in which its grammar allows different parts of speech — nouns, verbs, prepositions, etc — to be combined in "well-formed" sentences. Semantics is the other side of the coin; concerned with what the words in a language represent in the real world, and the way in which these individual meanings combine to convey the meaning of the string as a whole.

The central task of an NLP system is to "parse" these input strings — analyse them to identify their components and the relationships between them — using either syntactic or semantic strategies or, increasingly today, a combination of the two. But parsing is far from being the whole story. A complete natural language system also requires a dictionary or lexicon carrying all the vocabulary it needs to handle in the domain where it works, together with all the syntactic and semantic information about each word which it needs for processing — which may be very extensive. Some kind of translation module is required to convert the outputs of the parsing process into a useful form — commands to a database enquiry system, or structures from which text in a different natural language can be generated, for example. Systems handling

conversations or more complex enquiries will also have to have some deeper knowledge of the users' needs and the context in which they are working.

Clearly these capabilities add up to much more in combination than systems which simply take single words — spoken or keyed — and translate them directly into computer commands. Voice recognition technology in particular may provide a very important input to future NLP systems, but it must not be confused with them. This report draws a clear line between "single word" systems and those which are capable of some degree of parsing, by whatever technology, and so are rated as natural language systems. On the other hand, there is a huge range of capabilities between NL systems, and none of them come anywhere near the natural language ability of a human child — nor will they for the foreseeable future.

The value of natural language processing

There are two distinct types of application for NLP; to aid communication between humans and computers, and to aid communication between humans. The advantages of being able to use computers to handle natural languages are rather different in each case. As an interface to databases and other applications software, natural language is:

- understandable by everybody
- a universal, public, continuously updated standard
- able to provide a more flexible wide bandwidth interface to the underlying software
- almost always more concise than the same statement in a formal query language.

People are often sceptical about the concisenes of natural language, but compare:

> How do actual sales last month compare with forecast for people under quota in New England?

with the alternative equivalent:

```
PRINT NAME, 82-JUL-ACT-SALES, 82-JUL-EST-SALES,
82-JUL-ACT-SALES — 82-JUL-EST-SALES,
(82-JUL-ACT-SALES — 82-JUL-EST-SALES)/
82-JUL-ACT-SALES, IF REGION = 'NEW ENGLAND'
AND 82-YTD-ACT-SALES UNDER 82-QUOTA
```

As long as the system allows the user to add new vocabulary, it should virtually always be possible to phrase an NL statement which is shorter and easier to produce than the formal equivalent.

Against these points, there are some significant problems over the use of NL interfaces:

- a system has to be at least 90% successful in parsing queries from users to be acceptable, but even so there is still an element of uncertainty
- any system can be fooled; a human user can always find some expression which will deceive it, and this can be a good pretext for ignoring the system altogether
- interfaces may respond to ambiguous inputs by generating software commands which are not what the users wanted and not what they think they are getting
- NLP systems are usually expensive, and they take up considerable memory and processing power.

Problems like these have been one of the main factors holding back present-day NL interfaces to only a small share of the market. But looking into the future, the advantages seem bound to remain — or increase — while the disadvantages are steadily reduced. Each new generation of systems has wider linguistic coverage, so the proportion of failed queries will continue to decrease. The detection and resolution of ambiguity is improving in the same way. The falling cost of computer power will continue to improve the economics of NL interfaces. Driven by factors like these, the range of systems for which an NL interface is appropriate will grow steadily wider until eventually they become the rule rather than the exception.

The case for applying NLP to processing existing text is more straightforward. In applications of this kind, the computer system is either substituting for work which would otherwise be done manually, or else providing an extended service to carry out tasks which would previously have been left undone. So it may be justified because it saves labour, because it is faster or better, or because it can provide a service which would otherwise have been unavailable.

For example, there is already evidence that in favourable circumstances machine translation can double the productivity of translators — who are still needed to "post-edit" the computer output. It can help to improve the quality of translation by providing up-to-date, consistent equivalents of technical terms. And it can produce rough, unedited translations for information-scanning purposes which were often not available at all before. Similarly such applications as content scanning and the talkwriter will be able to offer direct productivity benefits to their users, combined with improved speed and accuracy in many cases. And again, the advantages must increase with time as the technology improves and computer performance increases.

Providing the technology

In the 1950s, some optimists thought that computer translation of languages would be quite quickly and easily achieved. They soon ran into difficulties when they tried to put their ideas into effect. The problem was not so much the occasional amusing word-for-word substitution — turning "hydraulic ram" into the Russian for "water sheep" must be the most-quoted example. This kind of error continues to crop up in MT, and probably always will; more fundamentally the machine translation pioneers found that language was more complex than they have allowed for, or than the hardware of the time could realistically process.

They found, for example, that the grammars which they programmed into these early systems would produce numerous alternative parses of the same sentence, without providing any way of distinguishing which one was intended. Consider an extreme example, like one quoted by Wendy Lehnert:

> "The appearance of the man under the tree with a broken branch near the edge of the road to the town with a market."

Human readers can make sense of the phrase because they have semantic knowledge. They know that the branch must be associated with the tree, the market refers to the town, and so on. But if the phrase referred to a broken arm instead of a broken branch, they would parse it differently, and subjected to purely syntactic analysis this one phrase is said to have no less than 429 different possible interpretations.

This is only the beginning of the problem. Besides the complications of regular grammar, a system operating in the real world must encounter a significant proportion of "ill-formed" input — meaning strings which do not conform with the rules of spelling and grammar the system has been given. Dealing with these invites a further combinatorial explosion in the number of possible interpretations.

Producing an effective talkwriter also demands the solution of much more complex problems than for a simple voice recognition system. Not only must it be able to recognise each phoneme — each distinct sound of the language — but it has to be able to assemble phonemes into words and words into sentences. Finding the boundaries between words in ordinary speech is very difficult, so most early talkwriter systems will require users to speak each word separately. Even if this is done, there is still ample scope for confusing words with similar sounds — and an appeal to the syntactic and semantic resources required by NLP will be essential to achieve a good level of performance. At the same time new hardware solutions will be needed to carry out the massive amount of processing required in real time.

With such a variety of problems to solve — and a corresponding variety of applications — it is not surprising that there is no single strategy or methodology which can be used to implement all NLP systems. Products on the market today use many different grammars, parsing techniques and ad hoc expedients, mostly dating back to the early 1970s. A new generation of database interfaces will be appearing in the next few years which will combine the syntactic and semantic approaches and have a sounder basis in linguistic theory. Leading examples are "Team", developed by SRI International, and "Irus" by Bolt Beranek and Newman. Other researchers are developing new grammars, such as GPSG (generalised phrase-structure grammar) and LFG (lexical functional grammar). Although it is still too early to say which will prove the most productive.

Other problems now being solved are to do more with the practicality and ease of use of NLP products. One of the main drawbacks to older MT and interface systems has been the time and linguistic expertise required to modify the system or add new vocabulary. New products, and updated versions of older ones, are now appearing which make this task much easier. Two products launched in 1985, Carnegie Group's "Language Craft" and Texas Instruments "NaturalLink Technology Package" go further, and provide tools which equip software engineers to build their own NL interfaces. Customising NL systems, and implementing simpler ones, is becoming a more routine matter.

As a result of these developments — and many others which there is no space to mention — the performance of natural language processing will improve. But NLP will never be trivial, and there are several major difficulties which are inherent in the nature of the task. For example, it takes a long time to develop and test NLP systems. People simply cannot master the complexity of language in a hurry, any more than they can learn a foreign language in a few weeks. Realistic user experience cannot be simulated or speeded up; it has to be accumulated over a period of time.

Secondly, the complexity of systems increases exponentially with their linguistic capability. To take one example, the Logos MT system uses about 20,000 rules to translate from German to English, and generates results which are about 80% the same as those produced by a human translator. Glenn Ransier, the product manager, estimates that the system would need 100,000 rules to produce results as good as a human translator — while still working on a limited domain of technical texts. Thus exponential growth in computer power and software capability can provide only a linear increase in NLP performance.

Taking these factors into account, the reader will understand why this report matches confidence in the ultimate importance and capability of NLP with a cautious view of its rate of progress. People looking for quick results from investing their time or money should not venture into this field, and this means it may have problems attracting the finance and staff numbers it needs. But the markets are there, the means to serve them are in view, and for those who have the stamina, natural language processing will prove a rewarding technology.

THE CURRENT SCENE

Emerging applications

One surprising conclusion of this study is how distinct and well-defined the applications of NLP prove to be. Language is so universal that the computer processing of language might be expected to have an undefined, all-pervasive role as well. In fact computers have such a limited capacity to cope with the full complexity of language that they can only perform a useful task if their work is focussed on a specific, restricted domain. In future it is likely that these application-specific systems will be linked together to provide more complex capabilities, such as the concept of a "corporate language system" discussed below. But for the present, each product or system is aimed at a specific task, and there is very little overlap or confusion between them.

Table 1 summarises the applications this study has identified, and the status of progress towards operational systems in each case. The table divides them into interfacing applications for human-computer communication, and text-processing applications to support human-to-human communications. They also divide into those which are already implemented as operational systems and products; those which are on the verge of operational use; and those where operational status looks a few years off.

Table 1 Seven major applications of natural language processing

Application	Definition	Status in mid-1985	Active companies & products
Database interfaces: mainframes and minis	Portable on integrated front-ends to DBMS and other applications software	Several commercial systems installed, starting in 1981; prices from $13,00 to $70,000	Artificial Intelligence Corp, (Intellect); Mathematica, (Ramis II English); Carnegie Group; IBS
Database interfaces: micros	PC or similar interfaces to application software or online services	First product in 1982; several in 1984 and 1985, with some 10,000's sold; prices from $75 (standalone) to £395 (package)	Excalibur (Savvy); Microrim (Clout); Microdata; Safeguard Business Systems; TI's Natural Link is an NL menu system
Dialogue interfaces	Conversational interface to complete system, eg expert system or ICAI	Prototypes include investment adviser, ICAI systems, none known to be operational	Cognitive Systems; several research programmes, including British Telecom, Logica, U of Cambridge in UK Alvey Programme
Content scanning	Processing messages or other semi-format texts to decide action or routing	A prototype for shipping messages due to be operational in 1985; systems for international banking telexes in progress	Cognitive Systems; Carnegie Group; Smart AI
Text editing	Checking style and grammar of text and suggesting improvements	At least one system undergoing extensive field trials by 1984	IBM (Epistle); AT&T's Writer's Workbench has some NLP features; other office systems cos. believed to be interested
Machine translation	Computer translation of texts from source to target natural language	Some mainframe systems have been in use since the 1960s; since 1983 mini and micro systems have been appearing	Systran Institute (mainframe systems); ALPS, Logos & Weidner are main suppliers of office systems
Talkwriter	Transcription of spoken input into computer text	Hardware under development; some promising demonstrations; product trials expected in 1986	IBM; Kurzweil; Speech Systems; major DARPA and Alvey research projects at MIT, CMU, Edinburgh

Mainframe interfaces

Table 1 shows that the applications which have already reached the market are mainframe and micro interfaces and machines translation. There is a big gap between the mainframe interfaces, typically costing tens of thousands of dollars, and the micro versions which can be as little as $75 in stand-alone form. Besides offering more breadth and depth of linguistic coverage, the mainframe systems provide much more in assistance and facilities for the user, support multiple users, and allow access to much more powerful DBMS and applications software.

But the progress of mainframe interfaces in the market has been relatively slow. For several years Artificial Intelligence Corporation (AIC) had the market almost to itself with "Intellect". Significant competition

appeared only in 1984, when Mathematica Products Group announced a natural language interface, called "English" integrated with its Ramis II mainframe DBMS and fourth-generation language (4GL) for database access. "English" does not seem to have done "Intellect" much damage. It stimulated one of Mathematica's main competitors, Information Builders Inc (IBI) to team up with AIC to produce a version of Intellect integrated with its own "Focus" DBMS, and AIC is now delivering more systems than ever before.

Other major DBMS suppliers, including IBM itself and Cullinane have already taken licenses to sell Intellect. IBM also has an advanced interface R&D project of its own, called TQA, which will probably be reflected in a product before long. Bolt Beranek and Newman, which has one of the biggest NLP research teams in any commercial company, plans to begin marketing its own NL interface development in 1986. By 1987 an NL interface should be a standard option for users of DBMS and "Information Centre" type software, and there will be a reasonable choice of alternatives.

So far, most NL products have been intended for IBM machines (or plug compatible, notably Amdahl). Two companies, Frey Associates and Intelligent Business Systems (IBS), are selling systems for DEC VAX users, although without much impact on the market yet. By early 1985 Carnegie Group, a new-venture company using the AI technology of Carnegie-Mellon University in Pittsburgh, was field testing a product called "Language Craft" which will allow software engineers to build their own NL interfaces for VAX systems. BBN also has a well-established interest in DEC markets, and its product could increase the choice for VAX users.

Micro interfaces

The micro market is completely different from that for mainframe NL interfaces, and mostly the preserve of small companies. Although mainstream NLP researchers are inclined to dismiss micro NL packages as trivial, and not offering true natural language, those already on the market are able to provide some interesting facilities, and customer response will provide a useful pointer to the demand for the new technology.

The pioneer in the field was "Savvy", developed by Excalibur Technologies in New Mexico. Savvy relies mainly on pattern-recognition techniques rather than parsing so it can accept quite a broad range of ill-formed inputs as long as they conform roughly to queries it already knows about. Microdata (now McDonnell Douglas Computer Systems) has developed an interface based on Savvy for DBMS access on its minicomputer line. Quite separately Safeguard Business Systems, a well-established business forms supplier, is selling an NL interface integrated with small-business accounting packages which has been developed by another small new company, Plain English Software Corporation, and this also appears to use a pattern-recognition approach. Sales of this type of system have been small so far, but one apparent advantage is its value as a demonstration tool — customers are attracted by the instant usability of the system even if they then choose to rely on conventional interfaces for real use.

Microrim has achieved more substantial sales with "Clout", originally an NL front-end for its RBase package for IBM PCs, and now offering access to a range of the more popular applications packages. Clout does parse the input string, allowing users greater freedom in the form of their queries, but without the robust acceptance of misspellings and other variants that Savvy can provide. Nominated as one of "PC Week's" 10 "Most Significant Products" for 1984, it has sold "tens of thousands" of copies by early 1985. By mid-1985 it was still the only known parsed natural language product available for micros, but forthcoming competition was expected from a PC version of AIC's Intellect, and from Symantec, a new-venture company with an information-management product in preparation.

The third strand of development in providing NL or near-NL access to micros is represented by Texas Instruments' "NaturalLink." This can be described as a menu system with a grammar. Users cannot express queries in their own words; instead they are presented with menus from which they select natural language phrases to build the query they want to express. The reason for claiming this as a natural language system, not just a very slick menu interface, is that the underlying grammar which determines which phrases users can select, and thus which commands they can give, has all the same functions and flexibility as that which underlies a "use your own words" system.

Aside from the issue of whether it is true natural language or not, there are strong arguments in favour of the NL menu approach, for example that it shows users the full range of things they can do, rather than leaving that to their imagination. Early in 1985, TI launched a "NaturalLink Technology Package" which equips software developers to write NL menu interfaces to their own products. It remains to be seen how popular this is; the NL menu could well be the most effective way of combining the limited resources of a micro with maximum ease of use. In the way of the package software market, it is quite likely that a product with some

kind of NL interface will catch the public's imagination sometime in the next year or two and spring into the best seller list to spawn a shoal of imitators.

Machine translation

The third application where NL products are established is machine translation. There is a substantial body of literature about the well-established mainframe MT systems, so instead of repeating that work this report focusses mainly on the results of two more recent trends which will be of great importance to commercial MT; the introduction of MT products which run on office minis and micros; and the move towards MT systems with a stronger theoretical base and a closer relationship with other NLP developments.

Systran Institute has been the main commercial supplier of mainframe MT, but three newer firms are leading its introduction to the office. Two of them, ALPS and Weidner, sprang from work in the remote state of Utah, inspired originally by the Mormon Church's desire to find a more efficient way to translate its gospels. The other, Logos, had almost opposite origins, translating weapons manuals needed for the Vietnam War. Now all three have identified their main market as being in Europe, where they had won a few score customers between them by mid-1985 — including, significantly, most of the major computer manufacturers. Smart AI, a less-known company, also has a system for translating pre-edited texts on a VAX.

The Europeans themselves have been behind in producing their own MT systems. Siemens has invested heavily in development in the USA, and ICL is participating in an English-Japanese project in the UK, but the most ambitious MT project in Europe is Eurotra which is financed by the Common Market, with the aim of building a modular multi-lingual system for the 1990s.

The Japanese have been much more active, Weidner now has a majority Japanese shareholding; Fujitsu, NEC, Sharp, IBM Japan and Nippon Telephone and Telegraph have all been working on mainframe or micro MT developments. But although Japanese work will obviously be very significant as far as the English-Japanese translation market is concerned, it is unlikely to have much relevance for translation between European languages in the near term.

Content scanning and text editing

The applications which can be described as on the verge of operational use are content scanning and text editing. Systems of both kinds have proved their worth in field trials, and should be going commercial during 1985 or 1986. Besides its Atrans product for reading bank telexes, Cognitive Systems has developed a system for the US Coastguard with a group at the State University of New York at Buffalo. This processes loosely formatted shipping messages and loads them to a database of worldwide shipping movements. It was reported as working very successfully as of late 1984. Carnegie Group is another company known to be working on a possible banking telex system.

IBM's "Epistle" is the leading text editing development. As a tool for helping writers of letters and other documents to identify and correct errors of grammar and style, it has had extensive field tests within IBM and with students. The barrier to making it a product may be more a matter of economics and computational efficiency than NLP performance, and it seems likely that both IBM and one or more competitors will be announcing products in 1985 or 1986. AT&T already sells a system called "Writer's Workbench" which carries out many of the same tasks but does very little language processing.

Dialogue interfaces and the talkwriter

Finally, the two applications which this study conclude are still a few years away from operational status are dialogue interfaces and the talkwriter. A dialogue interface goes beyond the simple query-response mode of a database front-end. It has to have some overall understanding of the user's plans and objectives, and to develop its model of his knowledge and changing needs as the conversation proceeds.

Useful dialogue interfaces are still at the research stage, and the underlying applications — which need to be clever enough to be worth having a dialogue with — are not yet fully developed either. Cognitive Systems is placing special development emphasis on "conversational advisor" systems, and has demonstrated one called "Le Courtier" which gives advice on investment strategies. BBN has an important research contract in the area, and in the UK British Telecom and the University of Cambridge are developing a system which will support a dialogue with telephone enquirers on topics such as train timetables. There are several other significant projects in progress, but it seems unlikely that true operational systems will be working before 1988 — and speech input will probably be essential for the most important systems.

This leaves until last the development which promises to be the most significant of all the applications of natural language processing — here called the "talkwriter." Even the name is still uncertain and controversial. Others in current use include "voice-activated word-processor" or "voice-activated typewriter", (which imply that speech is used just to turn the machine on), "speech-input word-processor" (equally cumbersome), and "VAT" (yet another mysterious acronym as far as the public is concerned). by comparison, "talkwriter" is a descriptive and convenient name, correctly suggesting a direct parallel with "typewriter." The usage goes back at least to August 1983, as a headline over a short piece in "Scientific American" (but not in the text itself); it has been used in a number of press articles since and it is favoured throughout this report.

Once successful talkwriters are developed they are expected to spread until they provide a routine working tool for millions of users who would rather speak than use a keyboard; an easy channel of communication with office systems and computers; and the central node in corporate NLP networks. The development costs will be high, but already there is a long list of organisations working towards this tempting goal. In the USA alone those known to be on this trail include IBM, two new venture companies, Kurzweil Applied Intelligence and Speech Systems, and the Massachusetts Institute of Technology and Carnegie-Mellon University, both funded by the Defence Advanced Research Projects Agency (DARPA). In Japan, NEC, Matsushita and Toshiba all have similar projects, and in the UK there is a major government-backed project led by Plessey.

The key issue is — how long will it be before one or more of these projects comes up with a useful product? The pace of development has already been more rapid than most experts would have dared to predict as recently as 1983. but in the spring of 1984 IBM demonstrated a research system with impressive performance in recognising complete sentences. At the same time both Kurzweil and Speech Systems were promising to deliver their first products in 1985, although now the official date for Speech Systems at least has slipped back to 1986. A reasonable prediction is that small test quantities will be shipped in 1986, but real products will not be available until 1987.

However, this will be only the beginning of the real test. With all previous major NLP products, years of user experience have been required to bring them to an operationally acceptable level. The talkwriter is likely to be at least as demanding. The early talkwriters will have severe limitations; users will have to train the machine by speaking a prepared text first; in most cases they will have to speak each word separately when they give dictation; the vocabulary accepted will be limited; the proportion of errors in the transcribed text may be high; nobody knows whether the provisions for coping with speech errors will be adequate; nobody has ever used machines like these, so completely unexpected human factors problems could arise when they arrive in offices.

The promoters of new-venture talkwriter companies are naturally inclined to gloss over these problems. No doubt they will be able to sell significant numbers of machines even in the early years because they will be bought for testing, to gain experience, and for specialised applications where they may be effective even in primitive form. But the prediction of this study is that the real take-off of the talkwriter into serving general-purpose markets, such as office dictation, will not come until well into the 1990s. When the takeoff does come, the talkwriter will become a major product and the core of a substantial industry. It could well be one of the symbolic products of the twenty-first century, as the motor car, the aeroplane, the television and the computer have been for the twentieth. But anyone who believes in some of the more absurd predictions, which show talkwriters reaching billion-dollar sales levels in the current decade, risks losing a lot of money.

Developments in the UK

Very few of the products and projects mentioned above are British. In fact, Britain is a long way behind the USA in the commercial applicationof NLP — if you can say that someone is behind in a race when they have not even started to run yet. This study found no British firm which is yet supplying an NLP product, and those few which are carrying out R&D in the area are mostly getting half their expenditure from the Alvey or Esprit Programmes — funding by the British Government and the European Commission respectively. A small number of mainframe interfaces and machine translation systems have been sold, and have won acceptance with users, but they have all been supplied by US firms.

On the other hand, British universities — including Cambridge, Edinburgh, Essex, Manchester and Sussex — have done some significant NLP research, particularly on the theoretical side, and are now beginning to work closely with industry on some key projects. A proposal from SRI International would further strengthen British NL research by establishing a specialist institute at Cambridge. These developments are stimulated by the £200m ($240m) of research funds which the British Government is providing through the Alvey Programme with a further £150m ($180m) from industry to finance research on "new-generation" computing in 1984-88. Together with the Esprit programme, this means there will be about £6.5m ($7.8m) a year available for NLP research in Britain. This is funding a small number of carefully selected projects which could include some of international significance.

Two major Alvey projects which have already been mentioned are the Plessey-led talkwriter development and the British Telecom telephone enquiry system, known as MAST (Machine Assisted Speech Transcription) and VODIS (Voice Operated Database Inquiry System) respectively. Other UK companies involved in Alvey NLP projects include ICL which has English to Japanese and Japanese to English MT projects with the Universities of Manchester and Sheffield; and Standard Telephones and Cables (the owner of ICL) which is working on another speech recognition project with Cambridge. Two systems houses working in the field are Scicon, which has an Esprit project, Loki, to build an NL database with the University of Hamburg, and Logica.

Finally, Eurotra should be included on any list of major NLP projects. Financed by the Common Market quite separately from Esprit, as of mid-1985 it was still purely an academic project, although discussions were beginning with systems houses for the design and development of the software to implement it. Eurotra could eventually form the basis of commercially important products if its objectives are realised. But it is less likely than projects such as Mast and Vodis to produce anything of operational value before the 1990s.

Strategies for using natural language processing

NLP technology already offers three different kinds of opportunity; using it, supplying it in conjunction with existing products, and investing in it by developing new products. Progress which has been made in implementing applications, and the benefits they have brought to users, are set out in the full report; the immediate user opportunities include the following:

- mainframe interfaces can generate significant added value by providing better access to corporate information resources for non-specialists; for example the fibreglass group of PPG Industries is using Intellect to provide close support to its marketing managers and sales people.
- micro interfaces are probably best seen as just an educational tool in NL for sophisticated users, but in many cases they are overcoming the barriers to computer use for neophytes.
- machine translation systems can already deliver productivity improvements, particularly for companies with large volumes of in-house technical translation — studies are given in the full report.
- content scanning could well be raising profits by cutting down transaction delays for some banks in the near future, and this is unlikely to be the only attractive application.

The experience of existing users shows that one vital requirement for the successful use of an NLP product at this stage is a committed internal champion. It is all too easy to make an NLP system look foolish, and a waste of time. Unless there is an advocate who is prepared to discover how the system can be used effectively, and educate people accordingly, it may quickly fall into disuse.

The next level of commitment to NLP is to enter the market as a third party supplier. This is a trend which is still in its infancy, but there are already some significant examples. IBM's decision to take a license to sell Intellect was a milestone, and as already discussed the product is now available in conjunction with a number of leading DBMS. IBI took a further step by working with AIC to produce Focus/Intellect which integrates the DBMS and the NL interface closely together. In another product area, Wang has worked with Logos to provide close integration between word-processing and machine translation.

Other manufacturers have used investments and research links to ensure that NL facilities are available on their products. A good example of such a symbiotic relationship is DEC's backing for Carnegie Group, and Carnegie's development of "Language Craft" which provides a tool for software engineers to produce their own NL interfaces for DEC machines.

Alternatively, tools like Language Craft and, for micro interfaces, Texas Instruments' NaturalLink Technology Package, will equip suppliers to develop NL interfaces to their own products themselves. If this is not feasible or desirable, there is an embryonic systems industry which has the capability to develop NLP software for third party use. For substantial NL packages, possible suppliers include BBN, Carnegie Group and Cognitive Systems. At the micro level, Safeguard Business Systems contracted with two academics to produce an NL accounting system and they have now incorporated themselves as Plain English Software. Microrim took a rather similar route for the development of Clout, and the NL experts involved have now formed Elliott Bay Computing to offer the same facilities.

Full commitment to NLP means direct investment in developing a new product or a new company. In fact there has been very little internal product development by established companies, but this report identifies 13 new ventures founded specifically to specialise in the natural language business. (They are ALPS, Artificial Intelligence Corporation, Cognitive Systems, Elliott Bay Computing, Excalibur Technologies, Intelligent Business Systems, Kurzweil Applied Intelligence, Logos, Plain English Software, Smart AI, Speech Systems, Symantec and Weidner Communications.) In all these cases, the long-term attractions have been sufficient to overcome the short-term difficulties of long product development and test times and uncertain pay-off.

But although nobody is yet known to have dropped out of the business, for a number of these companies it is questionable whether they have the resources to win through to an established, profitable existence — at least as they are currently set up. Some of them have been forced, by lack of long-term finance, to launch inadequate products and scramble for a few marginal sales to keep their heads above water. Others have been rescued from this position by timely injections of new investment. A suggested strategy for those intent on becoming a successful and substantial NLP supplier would go as follows.

1. Ensure that investment funds are committed for the long haul; this may downgrade the role of venture capital companies or even exclude them altogether if they stick to their standard objectives of a five-year payoff. The alternative source of funds is the major computer industry companies. Kurzweil has pointed the way by bringing in substantial investments from Wang and Xerox, both of whom have a long-term commitment to being major office systems suppliers.

2. Concentrate first on the R&D needed to bring the product's natural language performance up to the required level; this is what will take time, so the cost-effective computer implementation, and the product announcement, can largely wait until it is demonstrated. This seems to be more or less the strategy being followed by the established manufacturers doing NLP research — IBM, Hewlett-Packard, Xerox etc. When they do eventually announce products they will probably appear on the market quite quickly.

3. Draw on the research funding, and the wealth of academic ideas, available on both sides of the Atlantic. It may already be getting too late to take advantage of direct government funding for industrial research, from such initiatives as DARPA's Strategic Comuting programme and Alvey, but many academic projects are still not committed to any particular industrial sponsor. The head start they can give could make them a very profitable investment in a few years time.

4. As the moment for introducing a commercial product comes near, perform the trick of making a 180-degree turn from being a technology-based company to being a marketing-oriented one. Up to now, this has been one of the great weaknesses of NLP suppliers; an industry which depends on explaining its products and creating a felt want has been very weak on marketing. Some of the newer entrants, notably Symantec, seem to be investing much more energetically on the marketing side, and it will be interesting to see how effective this is.

The reward for implementing such a strategy successfully should eventually be a pole position in a major new industry. Natural language processing will be a strategically important technology for the whole office systems business, with a significant role in other areas such as aerospace and defence, which are not considered in depth here. Besides being vital to success in these areas, it should also be highly profitable to big suppliers, because of the barriers to the entry of competition — particularly the development lead time — and the substantial support and enhancement revenues from an existing user base. The reason for beginning work on the strategy now is that this market may never be so easy to enter again. The opportunities can now be identified; the resources and positions of strength needed to take them have not yet been pre-empted.

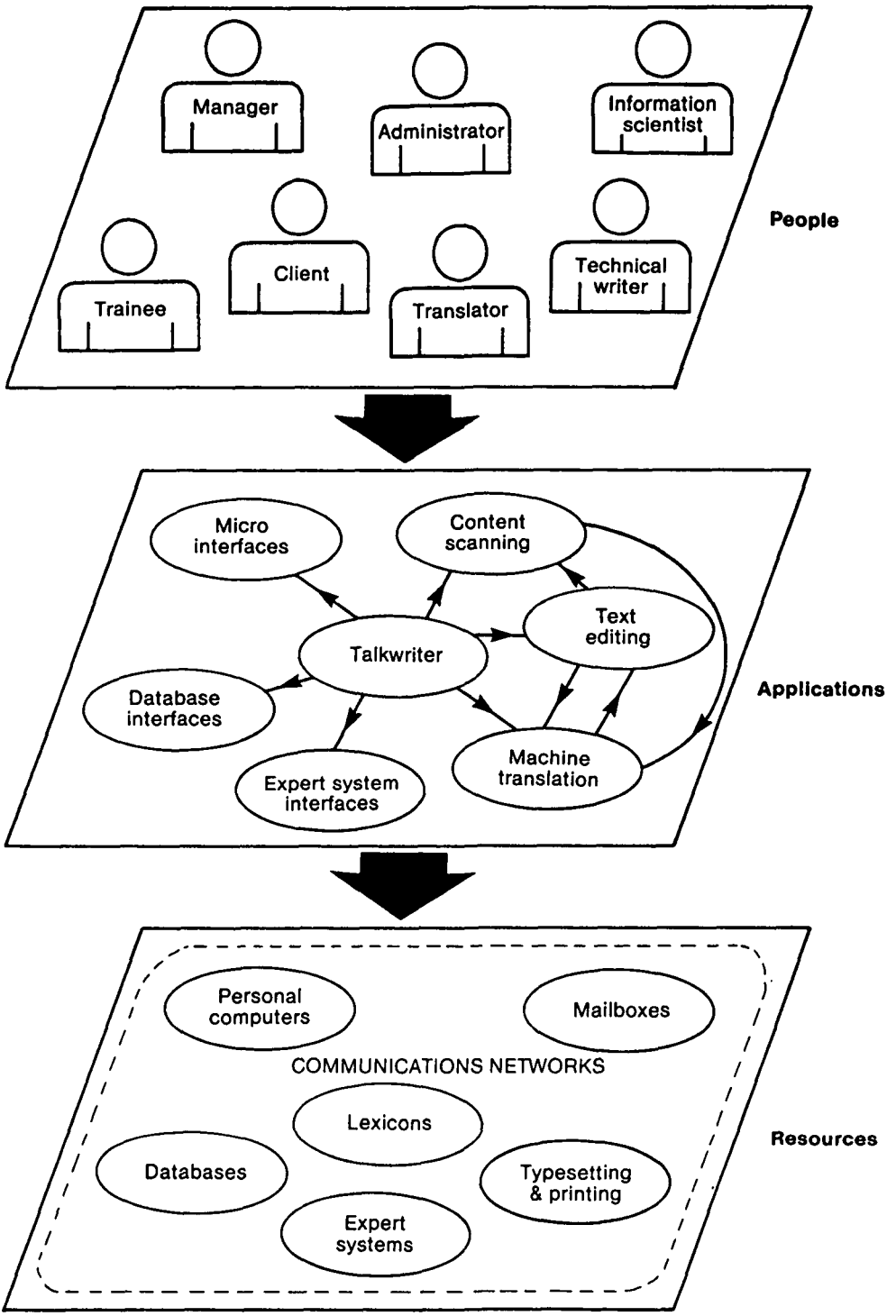
THE FUTURE OF NATURAL LANGUAGE PROCESSING

Towards corporate language systems

Most of the discussion until now has been concerned with NLP applications implemented independently of each other. But the different applications have a lot in common in the technology they use, the resources they need, and the people who communicate with them. In many cases separate applications could be linked together to provide more complete processing. Large organisations, and even medium-sized ones, could eventually have integrated natural language processing systems, along the lines illustrated in Fig 1. The long term future of NLP applications may lie along this route.

Fig 1 divides such a "corporate language system" into three levels — the people who use the system, the NLP applications with which they communicate, and the corporate resources which the applications use and to which they provide access. The diagram shows the most obvious potential links between applications — those between text editing, machine translation and content scanning are the only ones to be implemented at all at present — and indicates that all the resources would be closely tied together by digital communications networks. It does not show the multitude of potential links between the elements of the system at different levels.

Fig 1 The three levels of a corporate language system



The justification of such expensive systems will firstly be the need to avoid a new Babel (or at least to minimise the cacophony). As organisations gradually develop different, independent NLP applications, the problems of intercommunication will become increasingly significant. Even if the applications all nominally deal in the same natural language, each will have its own lexicon which will have its own vocabulary, semantics and language representation. A single common lexicon for all applications would be impossible, but an architecture which allows them to intercommunicate within the same framework *will* be essential.

More positively, the integration of applications should help to improve the overall performance of the organisation. In principle, an effective talkwriter system — backed up by machine translation if necessary — could give everyone in the organisation access to its information resources. Text editing and machine translation could assist anybody to produce fluent communications, across language and education barriers. Standardised text editing and lexicons will improve communication by ensuring that everybody uses the same vocabulary in the same way.

Finally, at the basic financial level, a CLS should offer significant resource-sharing benefits. Applications lexicons will represent a substantial investment, particularly with the accumulation of semantic knowledge and associations. The elapsed time required to create a lexicon from scratch may often be more important than the direct cost. The ability to draw on existing lexicons will be vital to the extension of applications within the organisation, and the development of new ones.

The social issues

Consideration of the functions of a CLS quickly brings more sinister issues to the surface — such as the possibility it offers of monitoring all information flows and controlling their content. Military intelligence agencies have had considerable interest in these aspects of NLP for many years, and there have been corresponding rumours, for example of systems which can identify key words in bugged telephone conversations. In theory a CLS could use microphones, talkwriters and content scanning to discover any "inappropriate" communications within the organisation, and integrated lexicons could make it linguistically impossible to express thoughts which were against corporate policy. George Orwell envisaged just such a requirement in "1984" when he invented Newspeak as a semantically and lexically restricted language in which it was possible to deal only in "goodthink".

More objectively, a CLS could be seen as a powerful tool which will help an organisation to carry out its existing policies more effectively. If it is already authoritarian and centralised, then the system will allow it to enforce more rigorous controls. If decision-making is distributed and decentralised, then the universal access made possible by a CLS will support an even more open style of operation. Even so, there will certainly be considerable concern about the implications of corporate language systems.

This concern is likely to be joined by increasing controversy about the role of NLP applications as a whole. The impact of MT has already aroused a bitter reaction by some translators; there could be a much louder outcry when talkwriters start to displace millions of typists and secretaries. The linguistic poverty of NLP systems, and the limited range of expression they force on their users, will certainly raise aesthetic objections. More generally, many nations will see their language as being much too important to leave to market forces, and may intervene to ensure that key NLP developments are under national control.

Establishing a system of this kind will be a major achievement — and a major investment. It is certainly not a facility which will be available in the near future. In fact, the scenario developed for this study suggests it will be close to 1995 before anything like the concept presented in Fig 1 is in operational use, even in leading edge organisations.

Similarly, the NLP applications may draw on different resources from the next level down. Content scanning, text editing and machine translation will be closely linked with mailboxes and with typesetting and printing. Micro interfaces will allow users to access a whole range of corporate resources through a personal computer. All the applications will depend on some sort of lexicon, in the sense of a computer dictionary, providing the definitive vocabulary for a particular domain. Each lexicon will be different, even employing different natural languages, but in an integrated corporate system, they will be structured in a complex hierarchy of their own which will, for example, allow definitions to be propagated in standardised form through all the applications, or restricted to a single domain.

Readers who try to trace some of the inter-level links will quickly discover that they are too numerous to show. A manager may communicate with a micro interface or a content scanning system. An administrator — anyone from a clerk to a CEO — may be working with database interfaces, content scanning or text editing. A client or customer from outside may communicate with a conversational expert system for assistance in purchasing the corporation's products, and so on. In all cases, the communication may be

direct, through a keyboard or a graphic interface, or — when such products become available — by voice, through a talkwriter system.

To add fuel to the flames, by the time natural language processing is a well-established business and growing rapidly there will be other, parallel, products of new-generation computer research reaching the market and having social impacts of their own. Image recognition, robotics and expert systems are among the technologies which together will change the pattern of employment across a wide range of professional, clerical and factory jobs. The Western economies — particularly the UK — are making a very poor job of redistributing the benefits of new technology across all classes of society. If the performance is not improved the reaction against some applications of the NLP could be very strong. It may yet prove more difficult to persuade people to talk to computers than to teach computers to understand them.