BEHAVIOUR & INFORMATION TECHNOLOGY, 1997, VOL. 16, NO. 3, 145–157

An analysis of the use of natural language processing systems in business

J. SIDHU† and C. J. HINDE††

†Department of Manufacturing Engineering, Nottingham Polytechnic, Burton Street, Nottingham, UK

††Department of Computer Studies, University of Technology, Loughborough, UK

Abstract. The results of a survey conducted amongst managers, users and application developers of Natural Language interrogation systems are presented and analysed. Those that were able to develop successful and effective applications using natural language paid careful attention to the certain stages. It is proposed that these stages are:

1. Systematic analysis of the company's requirements.
2. Effective integration of the natural language technology with the target database ensuring current applications are not adversely affected.
3. Introduction to new users of the system. This resulted in realistic user expectations and enabled effective use of the natural language software.

The advantages and disadvantages of natural language interfaces from an application developer, manager and user perspective are also discussed and recommendations made.

1. Introduction

The application of natural language has had mixed success (Manaris 1994: 228, Manaris *et al*. 1994: 260 and Sidhu 1991: 40). This paper presents the results of questionnaires distributed to companies who currently use natural language technology for database retrieval. The aims were to highlight any inadequacies of the systems and so information was obtained from people concerned with installing/maintaining, purchasing and using these systems. These groups were approached:

1. Application developers,
2. Senior managers,
3. Users.

Section 2 discusses the survey instruments and methodology, section 3 presents and discusses the results from the 3 questionnaires and section 4 contains the conclusions.

1.1. *Natural Language*

Bosser and Melchior (1986), Kalen and Allwood (1991: 81), Shneiderman (1993), McGraw (1994: 90) and Ramparany (1995: 388), argue that currently available database query systems do not meet users' requirements. Users can find it difficult to interact successfully with computer systems as the languages used are often unfamiliar and not suited to their problems; this is especially true of novice users (Shneiderman 1993). Many solutions have been presented to this problem (Sparck-Jones 1988:21) and commercial claims have been made that currently available natural language systems provide easier access to information from a database to the extent that they are able to provide reasonably good access to specific databases; access multiple remote databases; answer direct questions and do basic report generation (Hendrix 1982: 56).

The advantages and disadvantages of natural language are discussed in Shneiderman (1980), Hendrix (1982: 56), Vassiliou and Jarke (1984: 47), Morik (1984: 401), Rich (1985: 372), Ogden (1986: 51), Shneiderman (1986: 343), Whalen (1987: 1), Sidhu (1991: 32) and Sharit (1993: 32). In order to use a natural language database query system it was not just a case of buying the interface and posing queries in English; significant obstacles of varying degrees of difficulty had to be surmounted. Once overcome, these meant natural language technology could allow companies to expand their computing user base, leading to an increase in efficiency and ultimately productivity.

The questionnaire survey identified three stages involved in implementing a natural language system, and these are discussed in the following sections:

1.1.1. *Stage 1: systematic analysis of the company's requirements:* Failure to undertake this stage carefully

0144-929X/97 $12.00

may well mean total failure to meet objectives regardless of how well the subsequent stages are carried out. Sidhu (1991: 31) showed that companies paying meticulous attention to stage 1 had a better chance of making effective use of natural language systems. One company spent thousands of pounds on a natural language system which failed to deliver satisfactory performance in their environment. They claimed that natural language itself had not failed but they had failed to carry out stage 1 thoroughly. Engellieson and McBryde (1991: 43) report that users have difficulty with this stage because demonstration natural language systems are not available in their particular domain. They are unable to assess natural language as a tool in their particular application area. Even if a demonstration system is available, because no two-applications are identical the demonstration systems will not exactly reflect the intended application. Thus the purchaser must then decide whether their own application will achieve the same level of performance as the demonstration system. This is true of many software products but natural language is unique in the degree of software complexity involved which has a much more pronounced effect in stage 2 and the high level of user expectations addressed in stage 3.

1.1.2. *Stage 2: effective integration:* Integrating the natural language interface to a database has proved quite difficult (Hendrix 1982: 56, Obermeier 1990: 218 and Prechelt *et al*. 1993: 149). Some systems use such complex software that customizing the interface to the database requires an expert. Potential customers without direct access to such expertise may need to hire consultants to tailor the system, which can take time and is costly. As a result of hefty installation, customization and updating costs, many organizations may abandon natural language technology. Recently producers of natural language systems are offering do it yourself customization tools (Engellieson and McBryde 1991: 55). This, however, is conceptually very difficult (Hendrix and Walter 1987: 255, Prechelt *et al*. 1993: 149) and users are unable to use these tools effectively (Engellieson and McBryde 1991: 43). If this conceptually difficult process (stage 2) is not carried out comprehensively it will be very difficult to achieve acceptable performance levels.

1.1.3. *Stage 3: introduction of new users to the system:* Users and potential users have had their expectations of natural language systems raised to unrealistically high levels — Capindale and Crawford (1990: 341), Obermeier (1990: 218) and Bisantz and Sharit (1993: 32). The users of these systems, especially novice users, perceive natural language as the ability to communicate with a computer in everyday English. This is not the case, as the English used to communicate with databases via natural language is a subset of everyday English in a constrained domain — Suh and Perkins (1994:594). This is linked to the view that natural language is only significant when it is used to access a database of limited coverage. There is a trade-off between accessibility and the subset of English accepted i.e. a smaller database linked to a natural language interface is able to accept a wider range of English queries (Fano *et al*. 1989: 122 and Copestake and Sparck-Jones 1990: 248). Due to unrealistic expectations any interaction via natural language can leave users feeling frustrated and disappointed, particularly when questions which are input in perfect English are outside the subset of English acceptable to the computer — Capindale and Crawford (1990: 341). The natural language system is unable to interpret the correctly formed query thereby frustrating the users. Users need to be aware of the constraints and boundaries in order to use the interface efficiently.

2. Survey instruments and methodology

Nineteen organizations were approached of which 10 responded to the application development questionnaire, 12 to the management questionnaire and 12 to the user questionnaire which produced 19 respondents. Three products were included in the survey, Intellect by TRINZIC, Natural Language by Natural Language Inc. and STATUS /IQ by Harwell Computer Power. Table 1 is an overview of the organizations that responded to the survey and shows information on the type of organization; the significance and pervasiveness of the database; the complexity of the database domain; the complexity of the database in terms of the number of records and fields; whether the natural language application is a new application; whether new database operations are now being undertaken; whether new users are involved; whether an SQL is being replaced; the interface name; whether SQL is shown as output to the user and the response times from the natural language application.

2.1. *Evaluation*

By evaluating natural language systems currently being used in industry the impact of this technology can be assessed with regard to:

1. Acceptance by end users,
2. Comparison with the system it replaces,
3. The benefits.

Perhaps more importantly:

1. What users want improving,
2. Installation difficulties,
3. Overcoming these difficulties,
4. Users' overall opinion of natural language systems.

By conducting this survey the extent of the application of natural language querying systems in industry can be assessed. This work adopts the approach that a good computer system addresses the requirements of the real user, as opposed to the developers' perception of what users want.

Table 1. An overview of natural language applications surveyed.

Organization	Organization type	Significance pervasiveness of database	Database domain complexity	1. Database complexity 2. records 3. fields	1. New application 2. New database operation 3. new users	Replaced SQL	System name	SQL shown in query	Response time
1	Large International travel company	IT only	Complex	1. Complex 2. 500,000 3. <50	1. Yes 2. No 3. No	No	Intellect	Yes	Up to several minutes
2	Large service company	All departments	Complex	1. Complex 2. 100,000 3. 50	1. No 2. No No	No	Intellect	No	Up to several minutes
3	Large service company	All departments	Very complex	1. Complex 2. 4500000 3. >30	1. Yes 2. Yes 3. Yes	No	Intellect	No	Up to several hours
4	International energy company	IT only	Complex	1. Complex 2. 5 – 10,000 3. >30	1. No 2. Yes 3. Yes	Yes	Intellect	Yes	Up to several minutes
5	International energy company	All departments	Complex	1. Simple 2. 5 – 10,000 3. 50	1. No 2. No 3. Yes	Yes	Intellect	Yes	Up to several minutes
6	International transportation company	All departments	Very complex	1. Complex 2. 250000 3. >50	1. Yes 2. Yes 3. Yes	No	Intellect	No	Up to 3 minutes
7	Large health organization	IT only	Complex	1. Complex 2. 100,000 3. >40	1. No 2. No 3. No	Yes	Natural Language	No	Up to several hours
8	Large health organization	IT only	Very complex	1. Complex 2. 500,000 3. >50	1. Yes 2. Yes 3. Yes	No	Natural Language	No	Up to several hours
9	Large international food company	IT only	Complex	1. Complex 2. 500,000 3. >30	1. Yes 2. Yes 3. Yes	No	Natural Language	No	Up to several minutes
10	Large energy company	IT only	Complex	1. Complex 2. 160,000 3. >30	1. No 2. No 3. No	No	STATUS	No	Up to several minutes
11	Large county council	IT only	Simple	1. Simple 2. 50,000 3. 60	1. Yes 2. Yes 3. Yes	No	STATUS	No	Up to several minutes
12	Large international electronics company	IT only	Complex	1. Complex 2. 100,000 3. 50	1. No 2. No 3. No	No	STATUS	No	Up to several minutes

The survey involved assessment by questionnaires and in some cases an interview.

3. Results and discussion

3.1. *Application developers questionnaire*

This was aimed at application developers, the people that install and maintain systems, and who experience at first hand any technical difficulties and complaints of users. Ten questionnaires were collated.

This section paid particular attention to stage 2

3.1.1. *Results:* The results are shown in table 2.

3.1.1.1. *Stage 1: Requirements perspective:* Application developers felt immediate benefits of natural language included not having to train users in SQLs, increasing the user base by allowing familiar 'English' interaction, and users asking less of the information department. See appendix A for more opinions.

Of new natural language applications, 30% gave a marked productivity improvement. Of this 30%, 10% felt natural language gave a very marked improvement. Reasons for this included: inexperienced computer users finding interaction easier than they had imagined, users are able to ask questions without knowledge of computing (novice users), they are useful when used as an ad-hoc reporting system in a constrained knowledge domain, especially when the information is not readily available elsewhere as a report.

Of new applications, 50% resulted in no real productivity change and 10% actually felt they experienced reduced productivity. One reason was that the inclusion of natural language to a company's information system meant 'it took a long time to convert applications'. This can be linked to difficulties at stage 2.

3.1.1.2. *Stage 2: Technical perspective:* A major problem was the effective integration and customization required in order to use the system in the purchaser's particular domain. The majority (60%) felt that setting up natural language systems was easy or reasonable. Of this, 50% felt it was reasonably easy, and 10% felt it was easy; 40%, however, felt that setting up the natural language system was not easy.

Of the application developers 50% reported set up times to be in the order of weeks, 30% said it took days and 20% said it took them hours to set up. The time to set up in terms of man hours varied from 4 to 520, one installation taking as long as 3 months. The considerable variation in the set up time may be due to dealing

Table 2. Results of the application developer questionnaire.

Application developers
Natural Language is being assessed in any way by your company or organization.

Q1.) How easy was it to set up?

V. easy	0
Easy	10%
Reasonably	50%
Not easy	40%
Not easy at all	0

Q2.) What time period was involved?

Weeks	50%
Days	30%
Hours	20%
Less than 1 hour	0

Q3.) How long do you estimate it took to set up in terms of man hours?

Installation	Training
4 – 520 hrs	per user 2 – 16 hrs

Q4.) How easy do you feel the system is to use?

V. easy	10%
Quite easy	50%
Not easy	40%
Not easy at all	0%

Q5.) Do you feel casual/inexperienced users and experienced users are happy with the system?

Casual users		Experienced users	
V. happy	10%	V. happy	10%
Happy	30%	Happy	70%
Not happy	50%	Not happy	20%
Not at all happy	10%	Not at all happy	–

If users are not happy can you give any reasons?
Please use the reverse side if needed.

Q6.) How difficult is upgrading?

V. difficult	10%
Difficult	10%
Not a problem	80%

Q7.) What immediate benefits are there?

Q8.) In terms of productivity is there a favourable change as a result of utilizing a Natural Language system?

V. marked improvement	10%
Marked improvement	20%
No real change	50%
Negative improvement	10%

with very complex software systems and the difficulty of actually installing a complex system. It was noted that no one thought they were very easy to set up. Producers of natural language systems have attempted to sway potential purchasers of natural language systems with customization toolkits that enable easy installation and customization. This places the burden of installation on the users who can find this very difficult — Prechelt *et al.* (1993: 149) and Engellieson and McBryde (1991: 43). The times to set up varied from hours to weeks. This is surprising as talking to application developers in confidence left one with the impression that they were difficult to install and very difficult to customize:

> 'Adapting a natural language interface to a new application may be the most unappreciated problem in natural language design.'
>
> (Hendrox and Walter 1987: 254)

but in the questionnaire this was not borne out. Application developers may have felt uncomfortable as they are unique and traceable and hence have more to lose than users whose ability is not being assessed.

Of the application developers, 20% felt upgrading to be difficult, and 80% felt that it was not a problem. Thus upgrading was not a major problem after effective installation. This was not the impression obtained from informal discussions and interviews. Upgrading systems can be a difficult process (Sidhu 1991: 41). One application developer explained that after a difficult installation the natural language system was so imprecise that it could not be justified and as a result was slowly withdrawn from the company.

3.1.1.3. *Stage 3: Users' perspective:* Training varied from 2 hours to 3 days supporting the argument that natural language needs less training compared to SQL, as some people were expected to use systems after only 2 hours, though this was no guarantee of an initial successful interaction. From the companies surveyed the amount of training time, which varied from 2 hours to 16 hours, was not sufficient. Failure due to inadequate training and an initial unfavourable interaction by users can lead to users formulating negative impressions about natural language interfaces.

As a reflection of how easy the system was to use, 10% felt it was very easy, 50% felt it was quite easy and 40% felt it was not at all easy. Thus 40% of the applications surveyed did not achieve their full potential of allowing easy information access by users.

The survey showed that while 40% of casual inexperienced users and 80% of experienced users were happy with the system, 60% of casual inexperienced users and 20% of experienced users were unhappy with the system. From the tables it is clear that twice as many experienced as opposed to casual users were happy with the systems. These results contradict the idea that natural language is more suitable for casual inexperienced users (Bisantz and Sharit 1993: 32 and Morik 1984: 411 and Hendrix 1982: 57). This was further invalidated by the number of people who were unhappy with natural language: 60% of casual inexperienced users as opposed to 20% of the experienced users. One reason for this discontentment of casual inexperienced users was that they were not willing to learn how to use natural language systems because they had access to alternative products. Casual inexperienced users that were happy with the systems were those that felt comfortable with computers and had been carefully introduced to natural language. Especially so if users were able to access data they could not access before the introduction of a natural language system.

A frequent complaint about natural language systems was the claim that 'they can be used to communicate through English'. This caused users to harbour unrealistic expectations. This claim is partially correct but only a subset of English could be input. Other opinions are expressed in appendix B.

One viewpoint was that 'people began to associate the computer (natural language) as possessing real human intelligence, as this can not be the case the users may well be disappointed with natural language systems'. This lack of co-operative behaviour can be shown by considering the question

> Q) Were sales 10 years ago better than last year?
> A) No.

Is the answer to the question no, or is it no because insufficient data is held on the computer? The answer is not clear and shows lack of co-operative behaviour (Sparck-Jones 1988: 23).

A further criticism was that users had difficulty grasping the concept of natural language. Some questions produced 'two solutions to the same questions and one answer was not applicable'. The systems produced incorrect answers without the user being aware of this. In financial applications this was totally unacceptable, but this inconsistency only tended to occur with very complex queries.

3.2. *Management and user questionnaires*

This asked questions on aspects of stages 1 to 3 paying particular attention to stages 1 and 2. The management and user results are shown in Table 3 and 4 respectively.

3.2.1. *Results: Management questionnaire*

3.2.1.1. *Stage 1: Requirements perspective:* This involved assessing how well the objectives were achieved. Objectives in this context refer to whether the system is cost effective, the quantity of work produced is satisfactory, the productivity is good, and whether the overall level of performance is satisfactory. Of managers, 75% felt that the natural language system obtained was meeting its current objectives. Some users attributed the natural language systems' simplified syntax as one of the main reasons for this. The remaining 25% felt natural language was not meeting its current objectives. This may be explained by a failure to correctly match the currently available natural language technology with a company's requirements (stage 1). One senior manager was so impressed by a demonstration that he felt compelled to buy the system, only to have to stop supporting it as it didn't meet the required objectives. This emphasized the importance of carefully matching the company requirements with what natural language systems can achieve. Individuals, from the companies who failed in this regard, were able to link this to difficulties in integrating the systems to their databases (stage 2).

Of the managers, 33% felt natural language was cost effective, but the majority, 42% felt it was not. Dissatisfied managers felt that:

1. The systems are not being used as effectively as they could be. This is linked to failure at stage 1 or 2 or 3 or all three, as this would mean an inability to effectively achieve satisfactory performance levels.

Table 3. Results of the management questionnaire.

With the system are you satisfied:	Yes %	No %	N/A %
1. That the system meets its objectives?	75	25	–
2. That the system is cost effective?	33	42	25
3. With the quantity of work produced?	66	25	9
4. With the productivity of the system?	25	50	25
5. With its overall level of performance?	58	42	–
6. That the customers are getting a better product or service as a result of the natural language system?	58	42	–
7. That the users are getting the best out of the system?	25	75	–
8. That you are aware of any problems which exist with the system?	58	42	–
9. That you can resolve these problems?	33	50	17

2. Ineffective integration was problematic and costly and failure at stage 2, can be a major stumbling block. Natural language systems are complex software systems, so linking the interface onto a company database can be a difficult task, requiring total commitment. This work is often carried out internally, but if consultants have to be brought in it can be expensive.

The majority (66%) of managers surveyed were satisfied with the quantity of work produced but only 25% with the productivity of the system. This was contradicted by the majority (58%) who felt satisfied with the overall level of performance of their natural language system. Managers seemed happy with the amount of information available from a query but not the time and effort expended in obtaining it.

3.2.1.2. *Stage 2: Technical perspective:* The majority of managers (58%) were aware of problems that existed with the systems but only 1/3 felt that they could resolve them. Managers felt that:

1. Natural language requires a large time investment to set up. Installation and customization is often left up to the purchaser and can be difficult even with customization tools.
2. Significant software problems exist in current systems. No matter how well this stage is carried out, if there are significant software problems, as can occur with complex code, then the systems may never achieve acceptable performance levels. For further opinions see appendix C.

These problems apply to the majority of sophisticated software including SQL generators. Any problems will decrease as natural language systems advance and improve. This improvement will be accelerated as the number of companies producing natural language systems increases. The current market consists of isolated producers. If this isolation is broken down then areas of the market covered by natural language producers will overlap and they will be encouraged to improve their products in order to gain a competitive edge.

3.2.1.3. *Stage 3: Users' perspective:* The majority (58%) of management were satisfied that customers are getting a better product or service as a result of the natural language system. This is very encouraging from a potential purchaser's perspective. 75% however, felt that users were still not getting the best out of the system. This can be linked to stage 3. The smooth introduction of the systems to users is essential as initial unproductive interactions by users may sway them away from this technology. When assessing natural language

Table 4. Results of the user questionnaire.

	With the system are there problems:	No problems (%)	Minor problems (%)	Major problems (%)	Not sure
1.	Of poor documentation?	33	56	11	–
2.	Of obscure concepts which are difficult to learn?	22	66	12	–
3.	Of poor on-line help facilities?	47	42	11	–
4.	Of inconsistent ways of operation which are difficult to learn?	35	59	6	–
5.	Of losing control of the system?	42	47	11	–
6.	Of not knowing what to do next?	66	22	12	–
7.	Of having to remember too much information while carrying out important tasks?	47	47	6	–
8.	Of response times that are too slow just when you need a quick solution?	6	53	41	–
9.	Of response times that are unpredictable?	11	56	33	–
10.	Of occasional errors which are especially serious?	40	47	13	–

(a) Checklist on system commands.

		Yes (%)	No (%)	N/A
1.	Is a suitable prompt provided to make it clear what kind of response is required of the user at all times?	56	44	–
2.	Can commands be learned piecemeal, so that users do not have to learn a lot before they can do something useful?	100	–	–

(b) Checklist on documentation.

		Yes	No	N/A
1.	Can the user easily find relevant parts of the manual of other documents:			
	when getting started?	69	31	–
	when there is an error?	56	44	–
2.	Is the material comprehensible to novice users?	73	27	–
3.	Does the vocabulary correspond to that of the users?	62	38	–

(c) Checklist on being in control.

		Yes (%)	No (%)	N/A
1.	Does the system always behave in an expected manner?	16	84	–
2.	Are the menu options easily understandable, both in wording and layout?	85	15	–
3.	Is it possible for novice users to work in a safe limited part of the system, without needing to move into areas which are too difficult?	81	19	–

(d) Checklist on getting information out.

		Yes (%)	No (%)	N/A
1.	Is all the necessary information available on each display?	61	39	–
2.	Do items on the display appear in logical groups or sequences?	94	6	–
3.	Are all questions from the system clear?	56	44	–
4.	Are all question from the system necessary?	76	24	–

(e) Checklist on errors and error correction.

		Yes (%)	No (%)	N/A
1.	Are certain errors repeated frequently?	33	66	–
2.	Are experienced users prone to:			
	confusion errors?	37	63	–
	errors arising from inconsistencies within the system?	37	63	–
	errors arising from ambiguous instructions?	55	45	–
3.	Does the system help to prevent errors by requiring confirmation of inputs which may be ambiguous?	76	24	–
4.	Are error messages easily understood?	64	36	–
5.	Are error messages informative and helpful to the task in hand?	59	41	–
6.	Do error messages appear in a consistent manner and location?	83	17	–
7.	Is it possible to call for 'help' to supplement an error message?	44	56	–

it should be remembered that current retrieval systems, whether menus or SQL, are not in themselves totally satisfactory, hence any new alternatives do not have to be totally error free to compare favourably with current technology.

The survey shows that stage 3 shows room for improvement, i.e. the smooth and careful introduction of natural language systems so they can be utilized to a fuller extent. This level of dissatisfaction (75%) gives cause for concern as natural language is meant to simplify information access for people of varying computer experience — Bisantz and Sharit (1993: 32) and Morik (1984: 401). One manager expressed the opinion that: 'Specialised end user support skills are required, having a specialist who can be contacted when users experience problems may be just as important as initial training.' Further opinions are expressed in appendix B.

3.2.2. *Results: Users questionnaire:* This section focused on stage 3, the introduction of new users to the system and how well they were able to utilize the systems.

The effectiveness of training received by the users was assessed by asking questions about their ability to comprehend systems commands, documentation, whether they felt in control of the system, their ability to obtain information, being able to handle errors and error correction (Clegg *et al*. 1988).

Nineteen responses from 12 organizations were collated. The results are shown in table 4.

3.2.2.1. *Stage 3: User perspective:* Some companies did as little as two hours training per person and then users were expected to work with the systems using the associated documentation for further information. Of the users surveyed, 67% had some problems with the documentation, of these 11% had major problems. The survey showed users were generally satisfied with the documentation, 69% finding it comprehensible when getting started, 56% when there is an error, 58% when searching for cross references, and 73% when novices looked at the material. The survey showed an alarmingly high number (78%) had problems with the concepts of natural language, of this 12% encountered major problems; this was contrary to Hendrix (1982: 57).

Of the users, 53% had experienced problems with poor on-line help facilities, of these, 11% encountered major problems. Natural language systems were felt to be inconsistent in their operation by 65% of users and

goals
how does the artifact evoke goals in the user?
how does the artifact encourage users to import pre-existing goals?
how does the artifact suggest that a particular task goal is appropriate on inappropriate?
simple or difficult? basic or advanced? risky or safe?
planning
what distinctions must be understood in order to analyze a task goal into enabling methods:
how are these distinction conveyed by the artifact?
what planning mistakes are most likely; most costly?
how does the artifact encourage the use of background knowledge (concepts, metaphors, skills) in planning a task?
execution
how does the artifact make it easy or difficult to perform a task?
what slips are most likely? most costly?
how does the article indicate progress in task performance?
interpretation
what are the most salient features of the artifact? what do these features communicate to the user?
what features are commonly missed and at what cost?
what features of the artifact change as users carry out a task? what do these changes communicate to the user?
how does the artifact guide the user to make correct inferences?
what incorrect inference are most likely? most costly?
how does the artifact encourage the use of background knowledge in making inferences?
evaluation
how does the artifact convey completion of a task?
how does the artifact help users recognize and recover from errors?
how does the artifact encourage elaboration and retrieval of task goals and methods?

Figure 1. Posing questions about various stages of action can help identify the rationale implicit in usage scenarios (from Carroll and Rosson (1992)).

learning about these inconsistencies was a problem; of these 6% encountered major problems. Reasons included the same/similar questions which took varying times to produce answers, some of which are incorrect.

According to the survey 58% of users had difficulties with losing control of the systems, of these 11% felt this was a serious problem. The majority (76%) felt it was easy to move backwards in the system to regain control, having previously lost control. This earlier evidence contradicted the view that natural language enables the user to feel in control, and as a result instils confidence in the user. 34% of users found problems of knowing what to do next. This was especially so with users who felt, 'it can be difficult to see why a natural language sentence typed in is not understood.'

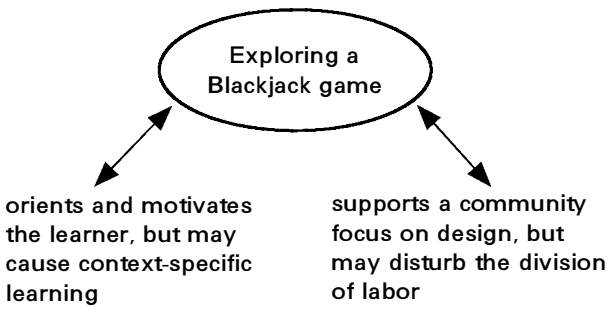

Figure 2. Scenarios of use can be developed at both the individual task level and at the organizational work level.

Stage 3 emphasizes how the concept of natural language falsely gives the user the impression they are in control. It is important that when training users on these systems that this false sense of security is checked. In the companies surveyed the majority of users were not made aware of this. Once the users felt comfortable with the concept of natural language they expressed positive opinions including that it is less intimidating and is more flexible as it does not require knowledge of an SQL. For further opinions see appendix A.

From the survey it was apparent that people associated natural language computing with spoken English. Not only is the English coverage of natural language systems limited it is not able to interact and ask for more information as in 'Sorry could you repeat that. I didn't quite understand what you said'. This coupled with casual inexperienced users' perceptions of computers as being very efficient and accurate lead people to have the view that computers are intelligent. In situations where inputs were rejected and users were left feeling confused because they thought they were interacting with an intelligent system, this is linked to high expectation. Users need to be guided on this

goals
how does the artifact evoke new goals in the organization?
how does it differentially evoke goals across personal roles in the organization?
how does the artifact support or conflict with various organizational goals?
how does it support the articulation of unspecified organizational goals?
planning
how does it support the coordination and conflict resolution among group members?
how does the artifact support the use of varying background knowledge in planning a task?
how does the artifact support the integration of individual planning efforts?
how does the artifact encourage vertical/horizontal line of communication in planning?
execution
how does the artifact support or require structuring work activities? (i.e., collaboration, division of labor)
how does the artifact limit or enlarge autonomy and scope of acvtion for various group members?
how does the artifact help the work group to coordinate and document progress?
interpretation
how does the artifact indicate progress in task performance to various group members?
what inferences and interpretations are most likely for persons in various organizational roles; how are they facilitated; what are their consequences for the group's learning and performance?
how does the artifact support the integration of individual interpretations?
how does the artifact encourage responsibility-taking?
evaluation
what kind of management structures does the artifact support, and how?
how does the artifact enhance or undermine group morale and social equilibrium?
how does the organization reward group members for successfully learning and using the technology?

Figure 3. Stages of action questions to help identify organizational claims in usage scenarios.

carefully and effectively, and is an essential component of stage 3.

Of the users, 94% felt slow response times were a problem, of these 41% felt this was a major problem. Users wanted some time response indication so they know how long it will take to answer an input. 89% of users felt response time unpredictability was a problem, of these 33% felt this was a major problem. Users views are included in appendix D.

A further criticism was that of occasional errors which are especially serious. 60% felt this was a problem and of these 13% felt it was a major problem. One company that dealt with financial transactions actually stopped using their natural language system because of the recurring problem of obtaining incorrect answers. This is an error inherent in complex software systems and it is difficult to explain to users during training without highlighting a very negative aspect of natural language. One approach is to encourage users to simplify complex queries as this may be less trying on the system.

3.2.2.2. *System commands:* After training to use natural language systems it was apparent that users felt reasonably comfortable with the system commands. 56% of users felt that a suitable prompt was provided making it clear when a response and the type of response that was required by the user. 100% of the users said they felt commands can be learned piecemeal, so that users do not have to learn a lot before they can do something useful. One committed organization constructed their own manual and small reference cards for a particular application. This organization taught users simple natural language commands in a familiar application. Users then progressed on to more complex commands. When suitable confidence was acquired users were encouraged to query outside this familiar application. This organization found that 'getting users to learn natural language concepts with an unfamiliar application is less than useless for the "average user"'. This is surprising as natural language is often portrayed as easy to learn, and an easy to use technology for information access to databases — Hendrix (1982: 57) and Bisantz and Sharit (1993: 32).

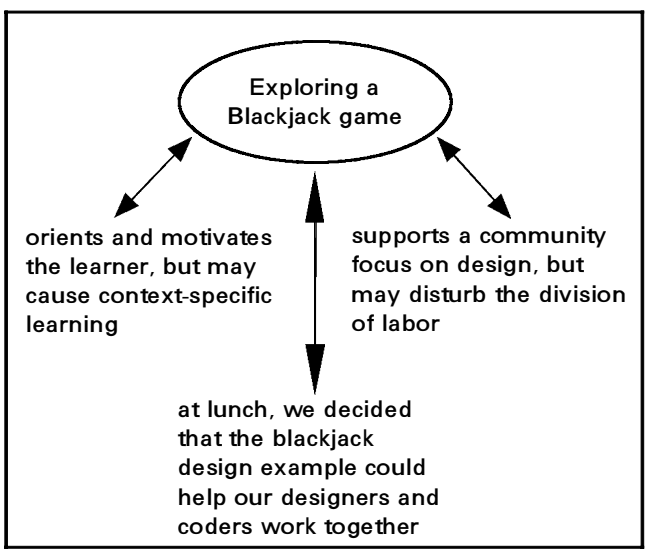

Figure 4. Scenarios of use can be explained in terms of the logical trade-offs in their rationale or in terms of the social contingencies that enable them.

3.2.2.3. *Documentation:* Users considered the natural language systems quite user friendly and their initial exposure was on the whole very positive. 69% thought easy references were possible to the manual/documents when getting started. 56% felt it was easy to locate the solutions to problems or errors. 73% felt the material was comprehensible to novice users and 62% that the vocabulary corresponded to that of the users.

3.2.2.4. *Being in control:* Overall it appeared users felt ill at ease with the wording and the ability to access information as a casual user. Users also felt the systems gave them problems in terms of erratic behaviour; some 84% of the users felt the system did not always behave in an expected manner.

A further problem was the possibility of obtaining an incorrect answer and not being aware of it. This is especially serious with applications that involve financial transactions. One company actually stopped using a natural language system because of the serious nature of this problem.

Of the users, 59% felt the 'on-line' help facility was adequate and 50% felt help information was presented clearly without interfering with the users work. Only 44% felt the help facility used language suitable for a novice user. A respectable 81% felt that it is possible for novice users to work in a safe limited part of the system without needing to move into areas which are too difficult. This is important as casual, inexperienced/novice users are able to work in their own domain safely having learnt just a few essential commands. This helps increase their confidence in the system.

3.2.2.5. *Getting information out:* 61% felt that all the necessary information was available on each display, and 76% that all the questions were necessary. 94% felt items were displayed in logical groups or sequences and 56% felt that all the questions from the system are clear.

3.2.2.6. *Errors and error correction:* Natural language systems are portrayed as being an easy to use technology (Hendrix 1982: 57, Morik 1984: 401, Hendrix and

Walter 1987: 251 and Obermeier 1990: 217). Questions were asked to assess how users were able to cope with making errors. This highlighted some serious shortfalls in natural language systems. 33% felt certain errors were repeated frequently, 37% of experienced users felt they were prone to confusion errors and 37% to errors due to inconsistencies with the system. 55% experienced errors arising from ambiguous instructions.

Of the users, 76% felt the systems tried to prevent errors by requiring confirmation of ambiguous inputs, 64% felt error messages were easily understood. 59% felt error messages were informative and helpful to the task in hand. 83% felt the error message appeared in a consistent manner and location, but only 44% were able to call for help to supplement an error message.

3.2.2.7. *Usability questionnaire:* Users felt the best aspects of natural language systems were the ease of use for technophobes, not needing a syntax, and the ability to ask questions in the way the user thinks. This is in agreement with Hendrix (1982: 57), Rich (1984: 40) and Napier *et al*. (1989). User opinions are in appendix E.

Problems included; 'the times it doesn't answer question as asked', poor response times, and inability to estimate response times prior to running enquiries. Depending on the complexity of the question the structuring of the query can be very time consuming because if the query is not precise then the output can be slow and incorrect. Two other problems were needing a good knowledge of the system to fully utilize its capabilities and the time taken to learn how to use the system in order to have confidence that the answers given are accurate and represent what was intended. For others please see appendix F.

3.3. *Discussion and recommendations*

Technical aspects of natural language which gave users difficulty included file linking. This is linked to the customization process in stage 2. 'File linking as an option, but corresponding response times so slow as to render it useless in all but the most basic application. A progress report which would indicate whether it was worth waiting for an answer or whether to abandon it and rephrase the query would be very useful'.

User-oriented aspects of natural language which gave users difficulty included the structuring and phrasing of expressions and reformatting questions that fail. This could be linked to a lack of training.

The users were asked about the changes they would like to see. Replies included: speeding up response times, being given an indication of the length of processing time, being able to type in requests for help in simple English, and the ability to speak to the machine rather than type. Speech technology is currently being developed but the ability to type in help and get appropriate commands will require an interactive machine and is beyond current capabilities. For further user comments see appendix G.

The impression of the natural language system from this survey was mixed. The technology does have the potential to be used effectively as a database query tool when certain stages (1 to 3) are carried out effectively.

Natural language has given greater functionality to the users. The support however, was seen to be lacking as some companies paid little attention to training (stage 3). The survey also showed that system designers were generally successful in moving the interface nearer to users thus expanding the user base.

If the access community does not include inexperienced or novice users then natural language may not be the best solution (Copestake and Sparck-Jones 1990: 225). Here it is proposed that integrating and customizing a natural language interface (stages 1 and 2) can be as costly as training users in structured query languages (stage 3).

The survey showed that natural language is good for companies in which there are large numbers of potential users with a good knowledge of the application. As one application stated:

> 'Natural Language systems are meant to be used by ''Intelligent people''. People who are familiar with the context of the database and know what tasks they have to perform.'

They tend however to be targeted for casual and inexperienced users.

Natural language systems have been used successfully for over 10 years and the introduction of customization tools has gone some way towards easing installation by dedicated staff but there is still room for improvement. This improvement will involve addressing issues such as users harbouring unrealistic expectations and improving system consistency. A further recommendation is to reduce variable response times and provide some feedback mechanism enabling users to have prior knowledge of how long queries will take. This will encourage users to rephrase complex time consuming queries into simpler queries. A major requirement of natural language systems is careful attention to training. The survey showed that training was not conducted as effectively as it needs to be. The careful control of users perceptions and expectations of natural language need to be redressed as preconceptions of this technology work against current systems. People perceive natural language today as being the ability to communicate with

computers using plain unrestricted English. The ability to do this is a considerable time away yet.

4. Conclusion

The application of natural language in industry was generally seen to be a favourable application if certain stages are completed efficiently and effectively.

Companies/organizations who paid careful attention and consideration to stages 1 to 3 were able to develop successful and effective applications.

1. Effective matching of currently available natural language technology to the company's requirements.
2. The effective integration of the natural language technology with the database, making sure the introduction of natural language technology does not adversely affect current applications.
3. A smooth introduction to new users of the system. This is essential firstly to enable effective access and use of natural language software, and secondly to control initial user perceptions which may be too high.

References

Bisantz, A. M. and Sharit, J. 1993, The effects of feedback on performance and retention of skill for a natural language interface, *Behaviour and Information Technology*, **12**(1), 23 – 47.

Bosser, T. and Melchior, E.-M. 1986, Learning in man-computer interaction: a critical review of the literature, ESPRIT project 385 HUFIT, report B3.6.

Capindale, R. A. and Crawford, R. G. 1990, Using a natural language interface with casual users, *International Journal of Man-Machine Studies*, **32,** 241 – 362.

Clegg, C. W., Warr, P., Green, T., Morik, A., Kemp, N., Allison, G. and Lansdale, M. 1988, People and Computers – How to Evaluate Your Company's New Technology (Ellis Horwood, Chichester).

Copestake, A. and Sparck-jones, K. 1990, Natural language interfaces to databases, *Knowledge Engineering Review*, **5,** 225 – 249.

Engellieson, B. and McByrde, R. 1991. *Natural Languages Markets: Commercial Strategies*, Ovum Ltd., ISBN 0903969610.

Fano, A. Kehler, A, Blejer, H and Flank, S. 1989. *Acquisition of Knowledge Sources for Natural Language Processing*, 122 – 129.

Hendrix, G. 1982, Natural language interface. *The American Journal of Computational Linguistics*, **1,** part 2, 56 – 61.

Hendrix, G. and Walter, B. A. 1987, The Intelligent Assistant. Byte, December 1987, 251 – 258.

Kalen, T. and Allwood, C. M. 1991, A survey of the training of computer users in Swedish companies, *Behaviour and Information Technology*, **10**(1), 81 – 90.

Manaris, B. Z. 1994, Natural language processing tools and environments: the field in perspective, in *Proceedings of the Sixth International Conference on Tools With AI*, 842.

Manaris, B. Z., Glanville, R. and Gillis, T. E. 1994, Developing natural language interfaces through NALIGE, 260 – 266.

McGraw, K. L. 1994, Knowledge acquisition and interface design, *IEEE Software*, **11**(6), 90 – 92.

Morik, K. 1984, Customers' requirements for natural language systems: results of an enquiry, *International Journal of Man Machine Studies*, **21,** 401 – 414.

Napier, H. A., Lane, D., Batsell, R. R. and Guadango, N. S. 1989, Impact of a restricted natural language interface on ease of learning and productivity, *Communications of the ACM*, **32,** 1190 – 1198.

Natural language, Natural Language Inc.

Obermeier, K. K. 1990, Natural selection, *Byte*, August, 217 – 222.

Ogden, W. C. 1986, Implications of a cognitive model of database query: comparison of a natural language, formal language and direct manipulation interface. In *SIGCHI Bulletin,* **18**(2), 51 – 54.

Prechelt, L., Finn, D. B. and Adams, R. 1993, Transportable language interface for taxonomic knowledge representation systems, in *Proceedings of the 9th Conference of AI for Applications* , 149 – 155.

Ramparany, F. 1995, Various uses of problem solving models for knowledge acquisition, *Proceedings of the 4th International Conference on Tools with Artificial Intelligence, Tools with AI* (Cat. no 92CH3203-7), 388 – 392.

Rich, E. 1984, Natural language interfaces, *Computer*, September, 39 – 47.

Rich, E. 1985, Natural language understanding: how natural can it be? *IEEE Conference On Artificial Intelligence Applications* , 372 – 377.

Shneiderman, B. 1980, *Software Psychology: Human Factors in Computer and Information Systems* (Little, Brown: Boston, MA).

Shneiderman, B. 1986, Seven plus or minus two central issues in human computer interaction, in *Human Factors In Computing Systems: CHI'86 Conference Proceedings*, M. Mantel, and Orbeton (eds), April 13 – 17, 343 – 249.

Shneiderman, B. 1993, *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (Addison-Wesley, Reading, MA).

Sidhu, J. 1991, Use of natural language in industry, Internal Report, Dept. of Manufacturing Engineering Nottingham Polytechnic.

Sparck-jones, K. 1988, Robust, co-operative and transportable natural language front ends to databases, *Computer, Science and Informatics, India*. **18,** 21 – 35.

Status iq, Harwell Computer Power Ltd.

Suh, S. K. and Perkins, W. C. 1994, The Effects of a System Echo in a Restricted Natural Language Database Interface For Novice Users, 594.

Trinzic Corporation, Intellect.

Vassilou, Y. and Jarke, M. 1994, Query languages-A taxonomy, in *Human Factors and Interactive Computer Systems*, Y. Vassiliou (ed.) (Ablex, Norwood NJ), 47 – 82.

Whalen, T. 1987, The feasibility of natural language interfaces for electronic database access, Technical memorandum, Behavioural Research, Department of Communications, April 1 – 14.