1987

TOWARDS UNDERSTANDING THE ROLE IMPORTANCE, AND IMPACT OF NATURAL LANGUAG INTERFACES TO DATABASES

Vijay Sethi
University of Pittsburgh

TOWARDS UNDERSTANDING THE ROLE, IMPORTANCE, AND IMPACT OF NATURAL LANGUAGE INTERFACES TO DATABASES

Vijay Sethi
Joseph M. Katz Graduate School of Business
University of Pittsburgh

ABSTRACT

This paper integrates diverse literature on natural language interfaces to databases, most of which emphasize technical issues. It formalizes one of the only categorizations of interfaces proposed in the literature. It also suggest how these systems may impact organizational activities, information systems usage, and decision makers' performance. A survey of information systems professionals regarding the use, role, importance, and perceptions about natural language interfaces is reported. The findings may be specially useful for developing implementation strategies for natural language based systems and further research in this area.

The introduction of artificial intelligence tools and applications has led to a growing body of literature analyzing their role in business. Natural Language (NL) processing is an important field that has many implications for information systems (Lyytinen 1985; Vassilou et al. 1983). However, because the area has been developed mainly by the disciplines of Computer Science, Psychology, and Computational Linguistics, inadequate attention has been paid to the managerial impact of NL based systems. Though studies providing state-of-the-art summarizations of different NL applications (Slocum 1985) present useful overviews to managers, they fail to describe the impact on decision makers and organizations. Currently, when NL based systems are moving out of research laboratories, information systems managers need to understand their benefits and evaluate them for organizational applications.

The focus of this paper is on NL interfaces to databases. Section 1 integrates previous literature and develops a comprehensive list of potential capabilities of NL interfaces. It also categorizes interfaces into four types depending on the extent to which they possess these features. Section 2 outlines some broad impacts of these systems and also explains how the effects vary with the type of interface. Section 3 briefly discusses some commercially available systems. Section 4 reports a survey in which information systems managers rated the relative importance of various NL applications, the importance of certain system capabilities, and the benefits of such interfaces. Following a discussion of the results in Section 5, highlights of the major findings and further research areas are outlined in Section 6.

1. INTEGRATING PREVIOUS LITERATURE

The first part of this section develops a comprehensive list of potential capabilities of NL interfaces based on past descriptions. The second part formalizes an existing typology of NL interfaces in terms of the previously derived capabilities. This enhanced categorization thus represents an integration of past literature.

1.1 Review of Past Studies

Various categorizations of the characteristics of NL interfaces have been proposed in the literature. Bates and Bobrow (1984) categorize interfaces according to coverage: a measure of the linguistic competence of a system, and habitability; the efficiency with which the user can recognize and adapt to the system's limitations. The dimensions used by Lehnert and Ringle (1982) are versatility: the range of functions a system performs, and discrimination; the degree to which users' actions and intentions are conformed with. Winograd (1983) identifies a number of design issues related to syntax, parsing, representation, meaning, and language. Bruce (1982) discusses that systems could exhibit knowledge about sentential form, semantic

form, discourse form, on-going form, and world knowledge.

Though previous studies use different terminologies, the system capabilities they describe are similar. Appendix 1 shows a comprehensive list of such capabilities that has been derived from the literature. Note that the capabilities refer to the extent to which systems can deal with different kinds of input only. Other features, such as transportability across domains or approach to ambiguity (multiple parses versus single parses), will be discussed as dependent characteristics because they are influenced by the previous twelve capabilities.

1.2 A Typology of NL Interfaces

One of the only classification of NL interfaces is by a panel of experts (Hendrix 1982), who propose three kinds of interfaces: Level 1 (L1), Level 2 (L2), and Level 3 (L3) systems. We expanded and formalized this typology by translating the interfaces' characteristics in terms of the twelve features described above and including the extensions of Level 1 system as a fourth type of interface. Because Hendrix describes the highlights of each type of system only, inferences had to be made to develop a comprehensive and uniform description of different interfaces. Table 1 describes the four types of interfaces in terms of the twelve capabilities.

The primary characteristic of L1 systems is that they incorporate an extremely limited theory of the application domain. The interface translates a query directly into a database command and does not maintain an explicit representation of the question.

Extended Level 1 (EL1) systems are more transportable across domains. They have limited inference capabilities because of functions that derive new information from primary data. These systems can handle pronouns and other references because they maintain the query context. Their response to

Table 1. Characteristics of Different Types of NL Interfaces

Capability	L1	E11	L2	L3
1. Grammaticality checking	low	moderate	high	high
2. Complex noun embeddings	rejected	mostly	mostly rejected	accepted accepted
3. Resolving pronouns	--	limited	moderate	high
4. Resolving other anaphoric references	--	--	moderate	high
5. Ellipses	mostly	mostly rejected	accepted accepted	accepted
6. Quantification	rejected	mostly	mostly rejected	accepted rejected
7. Inference	--	low	moderate	high
8. Understanding user goals	--	--	low	high
9. Knowledge of ordinary situations	--	--	low	high
10. Answering meta-questions	limited	low	high	high
11. Response to failure	pre-determined	----------"intelligent"--------		
12. Support in query composition	--	--	--	moderate

failure is "intelligent." They also give approximate responses whenever relevant, as well as answer some meta-questions.

L2 systems include an explicit theory of their domain of application; they incorporate internal representations of some domain objects. Translation is indirect, by converting to an intermediate form and then to a database query. These systems have extensive deductive capabilities, as well as provide discourse models for resolving anaphoric references. They can also handle telegraphic and ungrammatical input. However, they have limited abilities for managing quantifiers, time, and tense.

L3 systems maintain explicit theories about users, including information about their goals and plans. They translate queries indirectly by converting them to an intermediate form and using this information to make the following kinds of inferences: what the user meant; the type of information required to respond to the user; and the user's implied goals. These systems contain models of external situations and thus incorporate world knowledge.

2. IMPACT OF DIFFERENT TYPES OF INTERFACES

This section outlines in the form of propositions some broad impacts of NL interfaces and also describes how the effects will vary depending on the type of interface.

a. NL interfaces are likely to increase the number of information system users, especially top level managers. L3 systems will promote usage more than other types of interfaces.

Queries using NL, in contrast to those using a formal database language, do not require the use of artificial syntax or constructional elements (Cuff 1979). This is particularly helpful to a large set of people who are unwilling to learn or to use formal query languages. Completeness in a formal query language (emphasis on full syntax and precedence rule for combining logical operators) creates particular problems for casual users (Wallace 1984). The burden of learning an artificial language is also a severe barrier for top level managers who use computers infrequently. Even for technical specialists, learning many query languages when working on different systems can be difficult (Woods 1984).

Further, the use of NL interfaces does not require any substantial knowledge of the underlying database. Even users who are unfamiliar with the technical characteristics of a database can therefore formulate queries (Bates and Bobrow 1984).

Thus, NL interfaces are likely to promote usage because decision makers require less education and expertise as compared to formal query languages. However, different types of interfaces require varying degree of training. According to Rich (1985), the extent of training depends on the size of the sublanguage. Because L1 systems allow limited kinds of input, the size of the system sublanguage is small. At least moderate training is thus necessary for using L1 interfaces. By the same logic, L3 systems require almost no training. Training also depends on the extent to which users require knowledge of the underlying database. It is maximum for L1 systems because they provide a logical view, and minimum for L3 systems because they enable a conceptual view of the data (Harris 1985). Thus, because L3 systems are easiest to use, they would increase usage more than other kinds of interfaces.

b. NL interfaces are likely to increase the efficiency and effectiveness of users, and the improvement will be maximum for L3 systems.

Query formulation in a formal language may be much more complex than the corresponding English question (Wallace 1984). NL interfaces increase speed and convenience and decrease the likelihood of error because difficult queries can be formulated with ease (Petrick 1976). Further, interaction with the system can proceed in terms of what is needed without the specification of the retrieval method (Hendrix and Sacerdoti 1981).

Also, unlike computer programming languages, NL provides multiple levels of abstraction critical for solving business problems (Harris 1985). Thus, users would be more effective because NL provides a match between the user's conceptualization of the problem and the English structure appropriate for formulating the query naturally (Woods 1984).

However, L3 systems are likely to enhance efficiency and effectiveness more than other interfaces because only L3 interfaces can accommodate all user requests. In contrast, L1 and EL1 have very limited updating facilities. Updating is a problem because users phrase their queries with respect to their view of the database, which may be a simplification or a transformation of the actual structure. Update

requests may thus be "impossible" (cannot be performed), "ambiguous" (can be performed in several ways), or "pathological" (can be performed only in ways that cause undesirable side effects) (Davidson and Kaplan 1983). Handling update requests requires interfaces to understand the database structure and users' goals and intentions. Thus, only L3 interfaces would perform updating effectively. L1 systems would be effective primarily for retrieval, and L2 interfaces for retrieval and limited updating.

c. NL interfaces are likely to affect only certain types of organizational tasks. Further, the nature of tasks impacted by different interfaces would vary.

In general, NL is not a good medium in domains where analog rather than digital activities need to be described (Rich 1985). For such tasks, pointing, picture drawing, or steering wheel turning may be more suitable than language. Thus, text-editing is more efficient using an artificial command language than through a dialogue with a program. A match between the size of the language and the domain concept is therefore critical. Another condition under which NL interfaces are most useful is when the nature of the task is not well specified (Bates and Bobrow 1984).

The impact of different types of interfaces will differ on three major dimensions: task structure, task level, and ability to assist in problem solving.

On the dimension of task structure (Gorry and Scott Morton 1971), L1 interfaces would be useful only for structured decisions. This is because such systems handle limited kinds of queries only, a characteristic of routine tasks. In contrast, L3 systems provide access to multiple databases and can assist in ad-hoc, complex, non-routine query analysis. Thus, only L3 interfaces would be suitable for unstructured tasks. Similarly, EL1 systems would be suitable for relatively structured problems, and L2 interfaces for semi-structured problems.

On the classical dimensions of strategic planning, management control, and operational control, L1 interfaces would be suitable only for routine and operational tasks because of limited facilities for database query and access. L2 systems would be appropriate for management control activities because to some extent they can understand user goals. Because L3 systems are the most sophisticated, only they will be effective for planning and strategic decision making.

Interfaces also vary in their ability to assist in problem solving. This requires them to act as experts and possess the following kinds of abilities (Webber and Finin 1984): confirming or clarifying the user's understanding of the expert's advice; evaluating user suggested alternatives; justifying proposed solutions; clarifying doubts; and eliciting information from the user. Further, as discovered by Harry Tennant simulating a system (Cohen, Perrault and Allen 1982), suggesting alternatives to help users achieve their goals is also a critical feature. L3 systems would be most effective for problem solving tasks because of their sophisticated capabilities. Other interfaces would provide assistance in only some aspects of intelligence, design, and choice (Simon 1977).

d. NL interfaces may lead to user disappointment and frustration, and this threat is greatest for EL1 and L2 systems.

Users may be greatly disappointed if the benefits of NL interfaces are oversold or users are unaware of the limitations of these systems. While it is essential to propound the benefits of NL interfaces for initial use and ultimate acceptance of this technology, users should be aware that NL interfaces impose a number of restrictions on input, update requests, and so on. The threat of user disappointment is particularly grave for EL1 and L2 interfaces because they require users to understand that the limited inference and goal understanding capabilities of these systems do not imply that these interfaces are "smart," as is often believed (Hendrix 1982). On the other hand, users may find L1 systems too restrictive and the major problem of this interface would be restricting users to remain within the sub-language. Only L3 systems may be able to meet most user expectations.

3. CURRENTLY AVAILABLE COMMERCIAL NL INTERFACES

NL interfaces are fairly expensive to develop; Bates and Bobrow (1984) report that Cognitive Systems, Inc., of New Haven, Connecticut, is selling custom built interfaces for several hundred thousand dollars. However, less expensive interfaces are becoming increasingly available; existing products include Datatalker, from Natural Language Inc., costing about $10,000, Themis, which is available for

$1000 per node on a PC AT, and microcomputer interfaces from Symantec costing less than $500.

Current systems provide L1 and some capabilities of EL1 systems only. Specifically, existing NL interfaces enable reasonably good access to specific and multiple databases, answer direct questions, coordinate multiple files, handle simple use of pronouns, handle many elliptical inputs, enable basic report generation, analyze null answers (failed queries), correct spelling errors, and make limited inferences (Hendrix 1982). The INTELLECT system produced by the AI Corporation, Waltham, Massachusetts, is a typical L1 configuration. It is compatible with a number of database management systems on IBM, Honeywell, and Prime computers. Details of experimental systems, from ELIZA, SHRDLU to PLANES are provided in Waltz (1982).

4. IMPORTANCE AND DESIRED CAPABILITIES OF NL INTERFACES: A SURVEY

It was discussed that NL interfaces with limited capabilities are commercially available and researchers agree that "habitable, useful" interfaces are possible (Bates 1984). It is important for developing a strategy for the adoption of this technology to know the extent of their use and the perceptions of information systems managers about them.

Since NL interfaces are expensive, only large organizations would be able to afford them or consider acquiring them in the near future. Thus, a professional association called the "Pittsburgh Large User Group" consisting of organizations having extensive computing resources, including at least

Table 2. Profiles of Some Organizations and the Job Titles of Respondents

Organization Profiles:

1. A bank with domestic retail, worldwide commercial, trust and financial management services with assets of over $33 billion.

2. A manufacturer of polyurethane raw materials, plastics, chemicals, dyes and pigments with net sales of over $1.5 billion.

3. A provider of medical insurance to about 2.6 million subscribers.

4. A manufacturer of primary and fabricated aluminium and aluminium chemicals with sales of about $5.2 billion.

5. A manufacturer of consumer household products, toiletries, and proprietary medicine with sales of about $500 million.

6. A retail drug store with sales of about $652 million.

7. A steel maker employing about 50,000 people.

8. An explorer, producer, and distributor of natural gas with assets exceeding $3 billion.

9. A savings bank with total deposits exceeding $2 billion.

10. A manufacturer providing products and services to the construction industry.

Job Titles:

User Services Product Manager
General Manager of Computer Services
Vice President of Data Processing
EDP Research Consultant and Vice President
Systems Programmer/Analyst
Coordinator, Technical R&D
Project Leader (Special Projects)
Manager Advanced Design
Director, Data Administration and Security

one IBM mainframe, was contacted for acquiring data regarding different aspects of NL interfaces. A detailed questionnaire was mailed to the information systems/data processing managers of each of the 33 organization members. The president of the association made an announcement of the mailing at the group's monthly meeting.

The choice of subjects was dictated by the need to assess the importance of different aspects of NL applications to different organizations. Aggregating the responses of a broad cross-section of potential users was one option for obtaining this organizational data. However, because the number of users to survey was not clear and this was a preliminary study, it was decided to rely on the manager of information systems to assess the importance of NL based systems to their organizations. Nevertheless, a survey of actual users remains an important issue.
The response rate was 52% (n=17). This is fairly high considering a single mailing without follow-up. These results reflect a high interest of information systems managers in this new field. Though the sample is too small to make any broad generalizations, the findings could be very useful for future studies as little is known about the use and perceptions about this technology. The details of the questions and the results are described below.

4.1 Respondent Characteristics

Because of anonymity conditions imposed by the professional organization of the respondents, the survey did not elicit any company information. Organization profiles and the job titles of the subjects who volunteered such information are shown in Table 2. This representative data does not show any evidence of bias in terms of respondents' industries or positions.

4.2 Importance of NL Processing Applications

The respondents indicated on a five point scale, from "extremely useful" to "not useful at all," the extent to which eight NL applications (categorized in Winograd 1983) would be useful for their organizations. The answers were coded from zero to four, with four indicating extremely useful. Table 3 shows the average usefulness score of each application.

Respondents ranked NL interfaces to databases most useful for their organizations. Text retrieval systems ranked second, and text preparation systems ranked third. Machine translation was considered least useful.

4.3 Currently Used NL Interfaces

Only three (17.6%) of the seventeen organizations possessed NL interfaces. The installed systems were INTELLECT, CLOUT, and RAMIS ENGLISH DATA INQUIRY from Martin-Marietta. Out of the remaining fourteen organizations, two (12.5%) had considered acquiring an NL interface. The interfaces were planned to be used with a system for customer query for order status, and in aiding the operation of a large, complex data center. Both organizations planned to buy rather than develop in-house.

4.4 Importance of Capabilities of NL Interfaces

Respondents indicated the characteristics of NL interfaces that would be effective for their organizations by ranking different capabilities of NL interfaces. The results are shown in Table 4.

The top five desired capabilities of NL interfaces were efficiency, acceptance of pronoun references, transportability, resolution of elliptical input, and acceptance of sentences with complex noun embedding.

On an average, organizations required an NL interface with the following properties: moderate efficiency; providing multiple interpretations for most ambiguous sentences; limited transportability across domains; accepting some ill-formed sentences; mostly accepting complex noun embedding, complex noun phrases, and ill-formed input; resolve most pronoun references; and functioning without the use of any domain specific knowledge.

The average NL interface thus seems to be an EL1 system.

Individual responses were also analyzed for categorization into one of the four kinds of interfaces. We judged 25% (n=4) of desired systems to be L1 interfaces, 50% (n=8) to be EL1 interfaces, and 25% (n=4) to be L2 or L3 interfaces.

4.5 Perceptions about NL Interfaces

Respondents indicated their beliefs about the benefits of NL interfaces and the contingencies under which they thought systems would be most

Table 3. Applications of NL Processing and Their Perceived Importance

Application	Average Importance#
(1) Natural Language Interface to databases: These are front ends to databases and the system accepts queries in natural language.	2.88
(2) Text retrieval: Here the system interprets what information the user wants and retrieves the relevant document.	2.85
(3) Human-machine interaction: Systems could be used in situations where humans interact with machines using natural language not for question answering but for activities involving specifying what has to be done and monitoring what goes on.	2.73
(4) Aids to text preparation: Word processor capabilities can be enhanced by spelling checks and text critiquing facilities.	2.53
(5) Knowledge acquisition: Incorporating knowledge in programs, especially in programs called expert systems, can be tedious and costly. A natural language interaction between the program and an expert through which the program's knowledge could be built can be constructed.	2.44
(6) Computer Aided Instruction: Involves using computers for instruction purposes. Here the program "understands" user queries and answers them instead of typing pre-determined strings of words.	2.26
(7) Text analysis: Texts can be analyzed for use in intelligence gathering, for instance a program skims newspaper stories to keep track of the travels of a political figure.	1.82
(8) Machine translation: Involves translating texts from one language to another.	1.58

#: 0 = not useful at all; 1 = almost no use; 2 = somewhat useful; 3 = very useful; 4 = extremely useful

effective. Table 5 shows the results and details of the questions.

On an average, respondents mostly agreed that NL interfaces are very expensive, but somewhat agreed that they enhance user convenience. Subjects were unsure whether queries are answered with greater speed or if NL interfaces are easier to use than menu based systems.

Information systems managers consider NL interfaces to be most useful under the following conditions: when users do not understand the capabilities of the underlying database; when it is difficult to learn a formal query language; when the underlying database is unfriendly; when the task to be performed is somewhat unstructured; when the user's interaction with the system is not limited; and when users only query and do not update the database.

Eighty percent of respondents indicated lower level managers, 73.3% middle level, and 40% thought that top management would benefit most from NL interfaces.

Table 4. Average Perceived Importance of Capabilities/ Characteristics of NL Interfaces

Capabilities	Average Importance
1. Efficiency: 0 = not important at all 4 = extremely important	2.58
2. Pronoun references: 0 = reject all pronoun references 4 = resolve all pronoun references	2.44
3. Transportability across domains: 0 = general-purpose system 4 = special-purpose system	2.1
4. Ellipses: 0 = reject incomplete sentences 4 = fill omissions	1.82
5. Complex noun embeddings: 0 = reject all embedded clauses 4 = allow all embedded clauses	1.79
6. Parsing output: 0 = only semantic output 4 = only a syntactic parse tree	1.76
7. Response to ambiguity: 0 = only multiple interpretations 4 = only single interpretations	1.73
8. Complex noun phrases: 0 = reject all complex phrases 4 = allow all complex phrases	1.67
9. Grammaticality checking: 0 = accept all ill-formed sentences 4 = emphasize precise grammar	1.56
10. Domain Specificity: 0 = not at all, 4 = extensively take advantage of domain peculiarities.	1.21

5. DISCUSSION OF SURVEY RESULTS

Most organizations do not have NL processing systems. Information systems professionals consider NL interfaces and text retrieval systems more important than other applications of NL processing for their organizations. Machine translation is judged least important despite its long history of research (Slocum 1985). This may be because the sample did not include any multinational corporations where this application is proving very beneficial (International Management 1984).

NL interfaces desired by 75% of organizations were L1 or EL1 systems. Since L1 systems are commercially available and some capabilities of EL1 systems can be provided given the theoretical developments in this field, the demands of a large number of organizations can be met.

Some of the reasons why most organizations do not currently possess systems can be inferred from the perceptions about this technology. High cost of NL interfaces, as agreed by most respondents, may be a major barrier to acquiring such systems. Lack of knowledge or conviction about their benefits, especially as compared to menu based systems, might be another important reason. The perception that NL interfaces will impact lower and middle level managers much more than top level managers is surprising considering that most literature mentions their benefits to top executives only.

Table 5. Perceptions about NL Interfaces

Perceptions	Answers
0 = completely disagree; 1 = mostly disagree 2 = somewhat agree ; 3 = mostly agree 4 = completely agree	
1. Natural Language interfaces are easier to use than menu based interfaces.	2.06
2. Natural Language interfaces allow user queries to be answered with greater speed.	2.00
3. Natural Language interfaces enhance user convenience for queries:	2.59
4. Natural Language interfaces are generally very expensive and cost thousands of dollars.	3.12
5. Natural language interfaces will be most useful:	
(a) when the users do not know the capabilities or limitations of the underlying DBMS. (% agreeing)	76.5%
(b) when users cannot learn a formal interface language (% agreeing).	76.5%
(c) when the underlying interface is unfriendly. (% agreeing)	76.5%
(d) when the nature of the task to be performed by the user is: 0 = completely structured 4 = completely unstructured	2.59
(e) when the interactions of the users with the DBMS are: 0 = not limited at all 4 = extremely limited	2.32
(f) when the users are: 0 = experienced programmers 4 = extremely "naive"	2.50
(g) when they are used for: (% replying querying, % replying updating)	100%, 20%
(h) for the following level of managers: (% replying to top, middle, lower level managers)	40%, 73.3%, 80%

Maybe NL processing is associated more with efficiency than with effectiveness, and is therefore not considered very important.

6. CONCLUSION

Practical and usable NL interfaces are available today. Information systems professionals believe that they are the most important application of NL processing. Understanding the impact and extent of adoption of these systems is therefore an important issue.

Most organizations do not have NL interfaces. The survey suggests that this may be attributed to the high cost of systems and to a lack of knowledge or belief in their benefits. However, the capabilities of NL interfaces desired by most organizations can be met by existing systems.

The overview of NL interfaces provided in this paper may help practitioners evaluate the relevance of these systems for their organizations. Further, this paper can form the basis for more detailed studies. For instance, research can examine the extent to which implementation related factors, in addition to high cost and lack of awareness of benefits, inhibit the adoption of NL interfaces; establishing the appropriate syntactical and semantic connections between the interface and the database is a costly and time consuming process. Empirical verification of the impacts of these systems is another potential research area. Validation of the results of this study by expanding the sample size is also needed. Research can also analyze whether the perceptions of end-users regarding NL interfaces differs from those of information systems professionals. Thus, both researchers and practitioners may find this paper to be a useful first step towards understanding the potential role of these systems in organizations.

REFERENCES

Bates, M. "Accessing a Database with a Transportable Natural Language Interface." *Proceedings: The First Conference on Artificial Intelligence Applications*, Denver, December 5-7, 1984, 9-12.

Bates, M., and Bobrow, R. J. "Natural Language Interfaces: What's Here, What's Coming, and Who Needs It." In W. Reitmann (ed.), *Artificial Intelligence Applications for Business*, Ablex, New York, 1984.

Bruce, B. C. "Natural Communication between Person and Computer." In W. G. Lehnert and M. H. Ringle (eds.), *Strategies for Natural Language Processing*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1982, 55-88.

Cohen, P. R.; Perrault, C. R.; and Allen, J. F. "Beyond Question Answering." In W. G. Lehnert and M. H. Ringle (eds.), *Strategies for Natural Language Processing*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1982, 245-274.

Cuff, R. N. *Database Query Systems for the Casual User.* IBM Research Laboratory, Hursley, Winchester, Hants, 1979.

Davidson, J., and Kaplan, S. J. "Natural Language Access to Data Bases: Interpreting Update Requests." *American Journal of Computational Linguistics*, 19 (2), April-June 1983, 57-68.

Gorry, G. A. and Scott Morton, M. S. "A Framework for Management Information Systems." *Sloan Management Review*, 13 (1), Spring 1971, 55-70.

Harris, L. "Language for End Users." *Computerworld*, February 25, 1985.

Hendrix, G. G. "Natural Language Interfaces." *American Journal of Computational Linguistics*, 8(2), April-June, 1982, 56-61.

Hendrix, G. G., and Sacerdoti, E. D. "Natural Language Processing: The Field in Perspective." *BYTE*, September 1981, 304-352.

International Management "Translation Machines: The Smart New Tool for Multinationals." 39 (10), October 1984, 54-56.

Joshi, A. K.; Kaplan, J. S.; and Lee, R. M. "Approximate Responses from a Data Base Query System: An Application of Inferencing in Natural Language." *Proceedings: Fifth International Joint Conference on Artificial Intelligence*, Massachusetts Institute of Technology, Cambridge, MA, August 22-25, 1977, 211-212.

Krageloh, K., and Lockemann, P. C. "Access of Data Base Systems via Natural Language." In L. Bolc (ed.), *Natural Language Communication with Computers*, Springer-Verlag, Berlin Heidelberg, 1978, 49-86, (Lecture notes in Computer Science, #63).

Lehnert, W. G., and Ringle, M. H. (eds.). *Strategies for Natural Language Processing*. Lawrence Erlbaum Associates, Hillsdale, NJ, 1982.

Lyytinen, K. J. "Implications of Theories of Language for Information Systems." *MIS Quarterly*, March 1985, 61-74.

Malhotra, A. "Knowledge Based English Language Systems for Management Support: An Analysis of Requirements." *Fourth International Joint Conference on Artificial Intelligence*, September 1975, 842-847.

Mays, E. "Failures in Natural Language Systems: Applications to Data Base Query Systems." *Proceedings: The First Annual National Conference on Artificial Intelligence Applications*, Stanford University, August 18-20, 1980, 327-330.

McKeown, K. R. "Generating Relevant Explanations: Natural Language Responses to Questions about Database Structure." *Proceedings: The First Annual National Conference on Artificial Intelligence Applications*, Stanford University, August 18-20, 1980, 306-309.

Petrick, S. R. "On Natural Language Based Computer Systems." *IBM Journal of Research and Development*, July 1976, 314-325.

Rich, E. "Natural Language Understanding: How Natural Can It Be." *Proceedings: The Second Conference on Artificial Intelligence Applications*, Miami Beach, Florida, December 11-13, 1985.

Schank, R., and Abelson, R. *Scripts, Plans, Goals and Understanding*. Lawrence Erlbaum, Hillsdale, NJ, 1977.

Simon, H. A. *The New Science of Management Decisions*. Harper and Row, New York, 1977.

Slocum, J. "A Survey of Machine Translation: Its History, Current Status, and Future Prospects." *Computational Linguistics*, 11 (1), January-March 1985.

Thompson, C. W.; Ross, K. M.; Tennant, H. R.; and Saenz, R. M. "Building Usable Menu-Based Natural Language Interfaces to Databases." *Proceedings: Ninth International Conference on Very Large Databases*, Florence, Italy, October 31-November 2, 1983.

Vassilou, Y.; Jarke, M.; Stohr, E. A.; Turner, J. A.; and White, N. H. "Natural Language for Database Queries: A Laboratory Study." *MIS Quarterly*, December 1983, pp. 47-61.

Wallace, M. *Communicating with Databases in Natural Language*, Ellis Horwood Limited, Chicester, England, 1984.

Waltz, D. L. "The State of the Art in Natural Language Understanding." In W. G. Lehnert and M. H. Ringle (eds.), *Strategies for Natural Language Processing*, Lawrence Erlbaum Associates, Hillsdale, NJ, 1982.

Webber, B. L., and Finin, T. "In Response: Next Steps in Natural Language Interaction." In W. Reitmann (ed.), *Artificial Intelligence Applications for Business*, Ablex, New York, 1984, 179-194.

Winograd, T. *Language as a Cognitive Process*. Addison-Wesley, Reading, MA, 1983.

Woods, W. A. "Natural Language Communication with Machines: An Ongoing Goal." In W. Reitmann (ed.), *Artificial Intelligence Applications for Business*, Ablex, New York, 1984.

APPENDIX 1

POTENTIAL FEATURES OF NL INTERFACES

(1) Grammaticality checking: The system accepts ill formed sentences and does not emphasize precise grammar (Winograd 1984). This also enables "telegraphic input (input shortened by dropping unnecessary words)" (Bates and Bobrow 1984) and the use of abbreviations (Bates 1984).

(2) Complex noun embeddings: The system accepts complex, embedded sentences.

(3) Use of pronouns: The interface resolves pronouns. Pronouns usually refer to objects explicitly mentioned in previous discourse, but sometimes they can refer to objects mentioned later.

(4) Other references: The interface resolves other anaphoric references besides pronouns and also accepts multi-sentence utterances. Systems with discourse knowledge (Woods 1984) are one of the ways of resolving anaphoric references.

(5) Ellipses: The system fills in missing parts from context if portions of a sentence are left out (Woods 1984).

(6) Quantification: The interface interprets words like "some," "every," "all," and "any" using wide ranging "common-sense" knowledge or detailed knowledge of the particular domain (Bates and Bobrow 1984).

(7) Inference capability: The system draws logical conclusions based on domain knowledge and information in the database. This is necessary because a database rarely stores all required data explicitly.

(8) Understanding user goals: The system goes beyond the passive execution of users' commands and infers the goal structure underlying them (Woods 1984). It also volunteers additional information, recognizes and responds to users' misconceptions (Webber and Finin 1984), allows them to volunteer data, and gives approximate responses (Joshi, Kaplan and Lee 1977).

(9) Knowledge of ordinary situations: The interface has "world knowledge" (Lehnert and Ringle 1982). Scripts (Schank and Abelson 1977) are one of the techniques used to comprehend ordinary situations.

(10) Answering meta-questions: The system answers questions about the structure and organization of the database. Interfaces may adopt different approaches for this purpose (McKeown 1980).

(11) Response to failure: The system tailors its response to the type of failure, which may be intensional or extensional (Mays 1980), so as to provide maximum and correct information about its cause to the user.

(12) Support in query composition: The system provides help at query composition time. Menu based NL systems (Thompson et al. 1983) brought this support issue into focus and have been proposed as an effective means of providing help using "interaction experts."