

# THE AUTOMATIC CLASSIFICATION OF BUILDING MAINTENANCE

Douglas James Hague

PhD.

De Montfort University

Willmott Dixon Maintenance Ltd.,  
the Collaborating Organisation

September 1997

## Abstract

Abstracting reliable information from the operational levels of construction is a universal problem, and this research has focused upon building maintenance tasks of a maximum two days duration. Records provided by a private contractor were used in preference to local government or estate management sources, as being less likely to impose formal structures and thereby limit variety in the descriptions of repairs.

The subject of the thesis is the noise and disorder present in site feedback, and the ways in which they might be reduced to improve the quality of data.

The hypothesis was that by defining a minimum vocabulary, descriptions of repairs could be reduced through the elimination of unspecified terms, and from the remaining word-strings, make possible the automatic classification of maintenance.

A large sample of invoiced repair work carried out in the Midlands, was transcribed from original work sheets, provided by Willmott Dixon Maintenance Ltd., the collaborating organisation. Analyses of these data, which were characterized by their variety of tasks and unpredictability, displayed recurring patterns of syntax and word-phrases, which together with sentence reduction, encouraged support for the hypothesis.

An austere classification grammar proceeding from this examination is described. It proposes a lexicon of preferred building maintenance terms, two syntaxes; one using nouns hierarchically to denote headings of classes and levels, the other employing verb-phrases and adjectives for translating the original descriptions into reduced forms in a minimum vocabulary.

The conclusions are that there is sufficient support for the hypothesis to justify further research, this being largely to refine the linguistic theories and develop practical applications.

## Acknowledgements

I wish to acknowledge the assistance given by Willmott Dixon Maintenance Ltd. in providing the original information and for the trust shown by the Board of Directors allowing use of confidential and commercially sensitive material.

The advice and guidance of Dr. P. R. Innocent, Director of Study for this research, is also acknowledged, particularly in matters concerning Computing Science and natural language programming.

# Contents list.

	Page
Abstract	i
Acknowledgements	ii
Contents list	iii
List of illustrations	vii
Nomenclature	viii
1. Summary	1
2. Introduction	5
(a) Outline of the problem	6
(b) Aims and objectives	8
(c) Choice of a classification scheme	11
(d) The data	12
(e) The hypothesis	13
(f) Research method	14
3. Language of job descriptions	16
(a) Language system	17
(b) Structures of sentences in job descriptions	19
(c) Narrative descriptions	33
4. A minimum vocabulary	41
(a) Restricting the choice of words	42
(b) A definition of Minimum vocabulary	43
(c) The Construction Industry Thesaurus	44
(d) Consistency of word selection	46
(e) A strategy for selection of terms	47
(f) Part-of terms	48
(g) Type-of terms	50
(h) Type-of Building Elements	51
(j) Type-of Components	52
(k) Polysemy	53

(l) Cost as an attribute of selection	54
(m) Type-of material terms	55
(n) Summary and application of selection rules	56
5. A minimum vocabulary of actions and processes	62
(a) Conflicts of meaning and recurring phrases	63
(b) The place of meaning	65
(c) Experienced-word definitions	67
(d) Semantic diversity of <i>locating</i> and <i>clearing</i>	68
(e) Other uses for <i>remove</i>	74
(f) Verb to Noun-class with the <i>ing</i> suffix	77
(g) The prefix <i>Re</i>	79
(h) Specific descriptions	79
6. A review of classifications	82
(a) Classification approaches in general	83
(b) A review of some existing classifications for construction work	86
(c) Coding and maintenance classification	90
(d) Schedules of rates	95
7. A general classification structure for building maintenance identified and outlined	98
(a) Rationale for a new classification	99
(b) Arrangement of the proposed classification	100
(i) Building Zones	101
(ii) Services	102
(iii) External spaces	103
(iv) Scope	103
(c) Classification grammar	104
(d) Classification grammar compared with SMM7	106
(e) Syntax of 'Heading stage' grammar	109
(f) Verb-class patterns	111

(g) Vocabulary for the classification grammar	113
8. Maintenance data	118
(a) Data population	119
(b) Sampling method	119
(c) Confidentiality of records	120
(d) Identifying records	120
(e) Variability	122
(f) The archived material	123
(g) Inputting data by scanning	125
(h) Manual data input	127
9. Processing examined	130
(a) Computer programs	131
(b) Experiments with programs	132
(c) Criteria for an auto-classifier	139
(d) Concept of a developed system	139
(e) Organisation of an auto-classifier	141
(i) Read invoice description	141
(ii) Description for heading	141
(iii) Heading compiled?	143
(iv) Heading prefixed	143
(v) Job description parsed	143
(vi) Transformation matches heading?	143
(vii) Job, original description... formatted	143
10. Conclusions and recommendations	145
(a) Conclusion: the hypothesis	146
(b) Conclusion: the classification	147
(c) Recommendation	148

Appendix A Extracts from the CIT

I

Appendix B Correspondence

XVIII

Appendix C The classification of building maintenance	XXII
Appendix D Examples of original invoice records	XXVIII
Appendix E Computer program listings	XL
References	LIII
Bibliography	LVIII

## List of Figures.

	Page
Figure 3.1 Grammar, lexicon and phonetics (after Leech)	17
3.2 Sentence structures	20
3.3	22
3.4	23
3.5	24
3.6	26
3.7	27
3.8	28
3.9	29
3.10	30
3.11	32
4.1 Extract from Construction Industry Thesaurus	45
6.1 General classification scheme based upon Integrative Levels	86
6.2 Elemental Coding (after Holmes, Droop and Mellor, 1985)	93
7.1 Extract from proposed classification	101
7.2 Extract from SMM7 - F MASONRY	107
9.1 Screen output - Job No. 5631	135
9.2 Screen output - Job No. 5690	136
9.3 Example testing probability - first case	137
9.4 Example testing probability - second case	138
9.5 Comparison between normal approach to transformation and suggested method of auto-classification	140
9.6 Flowchart of automatic classification process	142
10.1 Research needs for the development of an automatic classification system	149



## Nomenclature

A	Adjective
ADV	Adverb
ADVP	Adverb-phrase
AI	Artificial intelligence
AP	Adjective-phrase
ASME	American Society of Mechanical Engineers
ATN	Augmented transition network
BASIC	Beginners all-purpose simplified instruction code
BEC	Building Employers Confederation
BRE	Building Research Establishment
BRS	Building Research Station (former title of BRE)
BS	British Standard
CAWS	Common arrangement of work sections
CCPI	Coordinating Committee for Project Information
CIB	International Council for Building Research Studies and Documentation
CI/SfB	UK version of international Samarbetskommitten fur Byggnadsfrigor (SfB) tables of classification
CIOB	Chartered Institute of Building
CIT	Thesaurus for the construction industry
COM	Comma
CONJ	Conjunction
D	Determiner
DP	Defining Participants
DOS	Disk operating system
IBM	International Business Machines
LAN	Local area network
M	Modifier
N	Noun
NBS	National Building Specification
NCR	No carbon-paper required
NLP	Natural language programming
NP	Noun-phrase
OCR	Optical character reader
P	Preposition

PP	Preposition-phrase
PRL	Princes Risborough Laboratory ( <i>vide.</i> BRE, BRS)
Prolog	PROgramming in LOGic: a fifth generation computer program language
PSA	Property Services Agency
RIBA	Royal Institute of British Architects
RICS	Royal Institution of Chartered Surveyors
RTN	Recursive transition network
S	Sentence
SMM7	Standard method of measurement of building works, 7th edition
UDC	Universal decimal classification.
UF	Used for ( <i>vide.</i> CIT)
V	Verb
VP	Verb-phrase

1.

Summary.

The principal objective of this research was to improve the efficiency and veracity in reporting building process data. Imprecision, irrelevance, expense and slowness in their collection, are failings inherent in almost every type and scale of project, and can have adverse effects upon operational management and supervision.

Building maintenance has several advantages for researching site communications.

Itemised details of the work done are necessary to claim payment, and documents in the form of orders, job dockets, time sheets and copy invoices, are collated and filed for accounting and auditing purposes. Most repairs are contracts of very short duration, embracing almost every attribute of larger and complex works, but without the formal practices of a site organisation. Data cannot be associated with a predictable construction sequence over time, making for variety in the type of tasks carried out. These last two factors, freedom for the individual manner of reporting work, and, absence of managerial decision over its choice, combine to make repetition virtually impossible. The research need for a large sample of records displaying characteristics of randomness, limited the enquiry to building companies, in preference to public sector estate management or direct labour departments run by large industrial concerns.

The research comprises of a primary investigation into the structure of language used to describe repair work, testing the hypothesis that limiting the vocabulary of descriptions will make it feasible to identify similarities for classification, and a secondary objective of proposing a possible application enabling automatic classification. Whilst computing science lies outside this present research, some aspects of computing have been investigated in order to present arguments supporting the concept, functional criteria and a speculative arrangement for a system.

The arrangement of this thesis broadly reflects the duality of the investigation.

Chapter 2 discusses the problems of obtaining reliable data about maintenance work, and defines the terms of research, its aims, objectives and the hypothesis. Methods currently adopted for codifying and scheduling building maintenance are reviewed

Analysis of the language used by the originators of invoices and work sheets begins in Chapter 3 and is further developed in Chapter 4, where the analysis of language is continued with reduction of sentences (transformation) and applying the principles of a minimum vocabulary. The effects of deleting proper names, prepositions and conjunctions with and various ambiguities examined, particularly arising from the role of noun-verbs. The *Thesaurus for the Construction Industry* (Roberts et. al. 1970) is introduced and its limitations as a vocabulary for maintenance are defined.

Verbs are further investigated in Chapter 5 and their role in relation to the object of a description analysed in depth. The proposition that classing descriptions by machine can be largely overcome by limiting the choice of words, is also developed.

Classifications and indexes are reviewed in Chapter 6, and those commonly used by the building industry, measured against a general classification based on principles of integrative levels. This Chapter marks the transition from a theoretical approach to making some proposals for a practical application.

The classification for building maintenance proposed in Chapter 7, departs from the numeric codes normally favoured for computerized methods, using instead words to denote categories and classes. A further divergence is that the classification is itself an austere grammar with two syntaxes, one of noun-phrases for defining categories, the other using in addition, verb-phrases and adjectives for translating descriptions from English into an English sub-set from a minimum vocabulary.

The documents provided by Willmott Dixon Ltd. the collaborating organisation are described in Chapter 8 and implications of source, and frequency of occurrence, upon their character are analysed and discussed.

The functional criteria, objectives and general outline of an automatic classification system is proposed in Chapter 9 Two computer languages are suggested for the software, QBASIC and Prolog, and their application is illustrated by means of examples of the output produced from program modules. The intention is to do no more than is necessary to support the feasibility of the collection and automatic classification of building maintenance by means of a system which, measured by the advanced computing science technology required, must be a tentative suggestion.

2.

Introduction.

(a) Outline of the problems.

Every building must be properly maintained to protect it from decay, whether an individually owned home or the most extensive public, corporate or private estate.

Distinctions between conservation, preservation, refurbishment and renovation, preventive, planned and periodic maintenance are unimportant; they are all repairs, made with traditional craft skills, adapted to new materials and methods. The simplest mend, whatever the context, is accomplished in an inevitable series of events.

First, a defect (or the symptoms of one) is discovered, enquiries are made and a potential method of diagnosis and correction identified. Next, an instruction is given and the repair carried out. Finally, a person (or legal entity) is charged for the work done.

There might be complicating factors, resulting in additional sequences, but in every case, the repair will be described, initially, by the operative who does the work. Such descriptions are expressed in terms and levels of sophistication, from lay to expert, subject to the originator's knowledge, training, expertise and perception. And it is upon this dubious base that managers build financial and information systems, regardless of their organisation's size and stature. Hobbs (1978) in her summary of a BRE Colloquium said.

"...input information is provided by the man who carries out the repair of a particular defect. He is required to record details about the work...It is important that this information is accurate because the quality of the entire system depends upon it. Indeed, some organisations consider that the uncertainty concerning the accuracy of the recorded data renders any detailed analysis invalid."(p. 2)

Establishing reliable data from the workplace is not a contemporary issue. Kelly, (1826) in pre-Victorian times, before master craftsmen had been displaced by building



contractors, appealed for improvements in submitted accounts:

"...when works are not to be measured...the only way the tradesman can make out his claim is to prefer a day account, specifying in general terms the works that have been performed...charged and compounded with the materials which have been consumed. To a considerate mind it would appear that objections could not be raised to this apparent and rational mode of delivering accounts... It is, therefore, to be regretted that our friends in the building profession do not, upon all occasions, keep and send in copies of their accounts in the most clear and intelligent manner, making distinct charges against EACH item; which system, if adopted, it is presumed, would prevent many, very unpleasant, disputes. (p. 47-48)

The following four records, taken *verbatim* from job sheets, not only give some idea of the last two writers' concerns for accuracy and clarity, they typify the kind of difficulty to be overcome when attempting to standardize operatives' descriptions of their work.

1. Squared up yard.
2. Adjusted spots to suit lamps lights okay.
3. Door is past repairing and is dragging on the ground.
4. Tested circuit with buzzer okay when manageress preferred.

Leaving aside linguistic analyses, the obvious comments are, (1) is ambiguous, (2) requires punctuation, possibly a comma placed between 'suit' and 'lamps', (3) does not describe a repair, it is a statement, and, (4) is open to several interpretations. They are the types of sentence structures which, potentially, can confound machine translation systems. More commonly, confusion arises with computers and English speakers alike, because many nouns also act as verbs. Given these grammatical dilemmas and the additional complications of misspellings, abbreviations, colloquialisms, trade terms and proprietary names, can job descriptions be comprehended automatically? The short answer must, of course, be 'No', since computers are incapable of constructing 'meaning' independently through reasoning; that is not to say they never will, neither does it preclude a solution. Resolving the current limitations of computers must be sought by

exploiting their data handling abilities and keeping strictly to the principle that information only has meaning when it reaffirms, creates new, or breaks old logical connections. What the human mind does to resolve such ambiguities as those in the example sentences, according to one paradigm (Miller, 1967), is an iterative process whereby inferences are drawn from the context of speech or writing, discarding unlikely and contradictory versions until an acceptable meaning is reached. Thus, When, and by Whom, information is *interpreted*, are variables of the event series previously described, not the translation itself

Summarising, the problem can be defined in these terms:

Any natural language produces, from even a limited vocabulary, several meaningful sentences, each capable of describing a single occurrence. Because the works are recorded in everyday English and technical terminology, the descriptions have to be translated into an austere dialect, or sub-set of English, according to some class of grammatical rules that will produce sentences, mutually exclusive for each type of transaction, however recounted. Therefore, the 'language' must become the classification to perform the transformations; it must have the rigour, and subtlety, to reveal salient identifying features and allow unambiguous allocation, without losing the capacity of being understood, when required, by a reader with technical building knowledge.

(b) Aims and objectives.

Given the complexity of the problems so far presented, the aim of this research can be expressed very simply:

to establish principles for the automatic classification of building maintenance work through the medium of naturally occurring job descriptions.

By 'automatic' is meant a self-acting process, whereby job descriptions are abstracted from an existing system and placed in the relevant categories of a classification without human intervention. Site operatives will not be required to use a particular phraseology,

or add allocation codes. Neither will the machine be programmed to interact with its operator, because an 'auto-classifier', by definition, has no need for reciprocity.

Therefore, it is envisaged that classifying descriptions would be carried out separately, even at another location, having been initiated by a routine invoicing procedure. This approach differs from what is usually understood as a computerized or computer-aided system, where effectiveness largely depends upon the quality of a dialogue with the user. More often than not, these interactions must conform to a protocol, involving training and lengthy periods of practice to become familiar with the software and understand how it is structured. Here, responsibility for the classification can be placed with managers and others qualified to determine whether job descriptions have been allocated correctly, and decide upon appropriate remedial action if they are not.

The precise nature and extent of these measures are dependent upon two components: how the principals of building maintenance firms would wish to use a classification of reported job data, and, the availability of a flexible, easily installed, usable, affordable system. Answers to the first are a function of the firm (with all that implies by way of structure and organisation) and will be as varied as the individuals concerned. But this diversity has a common imperative for intelligence on how the firm is operating. For example, financial systems were mentioned previously (para. (a)), and taking cost control as one component, Pilcher (1992) has emphasised the importance of having accurate knowledge;

'If a particular operation or process is being carried out inefficiently, immediate warning must be given to the ...management so that action can be taken forthwith to put the matter right.' (p. 403)

The variable nature and short duration of maintenance work add to the difficulty of pinpointing inadequate performance. Historical costing is not the problem being addressed, although details about repairs are normally reported after their completion; neither would everyone agree about the nature of cost nor which behaviour needs to be

controlled. Under investigation is the abstraction of data from a dynamic set of scattered events, where obtain verities of relevance, accuracy, immediacy and sufficiency. Most practitioners could recognise these as necessary, and concomitant, attributes applying to all operational information. Further, for descriptions to have meaning, they must be associated with a context and grouped, like with like. The first objective is then:

to compile a general classification for differentiating between types of building maintenance work and provide an adaptable framework for various management systems.

Predicting the availability of an automatic classifying system, the second component restraining managerial responses, must of necessity, be based upon informed opinion and is, to some extent, speculative. Developments in micro-computing, both of hardware (machines and devices) and software (coded instructions to organise functions), present a landscape of accelerating change and technological advance. What was either barely feasible or uneconomic, can very quickly become commonplace and relatively inexpensive. There is a strong and highly competitive international market, and in Britain, specialist software companies and individual entrepreneurs write powerful and innovative applications for the construction industry. Development of a marketable system lies entirely with computing scientists, and therefore, cannot be an avenue for this present research to explore. Nevertheless, the requirements of an auto-classifier have to be described in sufficient detail for technological and commercial evaluation.

Consequently, the second objective is:

to define the parameters within which an automatic system capable of abstracting, processing and presenting a classification of building maintenance data can be described.

(c) Choice of a classification scheme.

Assuming it is feasible to achieve the stated objectives, what would be the best method of classification?

The public sector, with the predominate investment in building maintenance, offered a credible source for this 'best' scheme. Seeking to improve the costing of management and maintenance, initially in local authority housing, then school buildings, coding and cost classifications have been established to allow monitoring and analysis of feedback information (Holmes and Mellor, 1985). The generally adopted approach, as described by Holmes, Droop and Mellor (1985) is an heirarchical code; they do not report any difficulty in recording data 'even as many as 15 digits is not as laborious in practice as it might appear...'. Then a question posed by Skinner and Kroll (1984) has to be considered - the method and detail of coding;

'The choice of coding method is... an important decision because it governs the amount of future feedback which can be obtained, possibly 20 or more years ahead; an interim change could effectively waste the preceding maintenance data unless experience with the original coding supports a reduction of detail which is compatible with the original system.' (p. 54)

Similar considerations, on a reduced timescale, must apply here if elements could be defined for the range of descriptions yet to be translated.

Hand-in-hand with effective monitoring, the management of work has become more rigorous. For at least 30 years, successive governments have encouraged the use of measured term contracts, with their concomitant schedules of rates to provide the basis for tender assessment and contractor payment (or crediting direct labour departments). Broadly, schedules are of two types; **detailed**, containing individual items of work or activities, and, **composite** with items comprising several, associated or consecutive, activities. All items have identifying codes, usually following the Standard Method

Measurement of Building Works (7th edition) system of classification (RICS/BEC, 1988). Reports on operating systems by Saville (1990) and Holmes (1987), describe them as fully computerised, the coding and items being transcribed by clerical staff, directly supervised by maintenance managers. The task of writing job descriptions at the work-place is rendered unnecessary, as Holmes, Droop and Mellor (1985) claim:

'(To make the system foolproof the final verbal description of the code that has been typed would be displayed on the screen as a check). Moreover, the person 'phoning in is unlikely to know the detail of the operation so the level of coding required at that point is limited.' (p 10)

According to anecdotal evidence, operatives keep schedules of rates in their vans for reference when compiling (then apparently, telephoning) reports, thus enabling them to quote relevant codes. Clearly, this practice is encouraged because not only does it conform to the *raison d'etre* for adopting schedules and measured term contracts, the non-technical staff are further helped in their routine analysis and processing of contractors accounts. Whether the veracity of the data is preserved is questionable; an intelligent workman could select codes and items to give a chargeable time and value for the repair greater than the actual time spent on site. (Gold, 1990 cites more serious frauds). In this sense, items and codes choose descriptions, whereas this present research intends to reverse the order of selection. Taken together with an uncertain outcome of restricting the number of classifying hierarchical elements or using them all, with the alternative of matching descriptions to any one of the published schedules, adopting any form of predetermined listing would be incompatible with the research objectives.

(d) The data.

Seeking a method to classify repair works automatically, requires original descriptions of the repairs, and not paraphrases from a schedule or specification. Fortunately, Willmott Dixon Maintenance Ltd. agreed to collaborate in this project, and generously provided

data. These were copies of sets of documents used for the preparation of invoices, and nearly 2000 were examined and the relevant material extracted from them. Responsibility for this commercially sensitive information was not lightly accepted, and every care was taken to keep it as secure as if in the owner's archive.

(e) The hypothesis.

The need for information about building repairs and the difficulties in obtaining it, was first brought to the attention of the present author in 1966, during a period of secondment to the Construction Industry Training Board. Organising courses for the principals of small building firms, gave a unique opportunity to discuss the problems of managing repair work, in which the majority of the course members were engaged in. Over the succeeding years the general problems of small firms have been considered and some resolved, but the transaction of data between site and office has remained a barrier to fully utilising information technology. As discussed above, this basic organisational weakness is typified by tradesmen, untrained in clerical skills, providing details of their work for clerical staff who use the data for processes vital to the firm's success. A focus upon the issues involved, began with unpublished research (Hague,1983). Evidence seemed to support a view that maintenance tasks were limited in range, if not tending to be repetitive. Variations were often attributable to the context in which they were carried out, the language used in their description, or both factors together. Although sample sizes of job durations were small (less than fifty for any contributing source), distributions bore a strong resemblance to the positively skewed shape reported by Allan (1966) and Kirby (1970). Some repairs were more common than others, blocked drains; slipped/broken roof tiles; leaks in flat roofs; easing doors. Complexity, when it appeared, lay in the description, rather than the content of the task itself. Reasons advanced by

principals for style and amount of detail presented in an invoice, were functions of their clients' requirements, or even to justify what might be seen as high charges in relation to the repairs carried out.

An earlier attempt to capture job descriptions electronically, (Hague, 1983, pp. 6.29-32) provided the concept for this present research. The experiment sought to establish similarities between reported repairs by identifying key words with bar codes, then compiling descriptions from a list with a hand-held reader (such as those used in shops and stores), and allocate them to a schedule or database. Given the advances in computer technology since the early 1980's, it now seems worthwhile reopening this line of enquiry, taking a theoretical approach to examine the nature of the language used in reporting repairs. The practical background to this research is expressed in the stated aims and objectives, their main purpose being to give points of reference for the following hypothesis: :

that by limiting the terms used in written descriptions of building maintenance, the resulting abbreviated word-strings would enable repairs to be labelled and classified automatically.

Evidence to justify, or refute, this contention was tested against the *possibility* of developing some workable system, not the development itself.

(f) Research method.

Enquiries began with a literature review of building maintenance management, classification theory, linguistics, applied language studies and natural language programming. Bibliographic and reference listings were provided by the Chartered Institute of Building (CIOB) from its own database, and De Montfort University's Kimberlin Library, whose search covered several sources including the RIBA Database, INSPEC, and Compudex. These were supplemented with Building Research



Establishment publications; Current Papers and Building Digests, CIOB publications; Site Managers Information Service, Technical Information Service and Construction Papers, in the present author's private collection.

The main plank of enquiry was an examination of the copy invoices and documents provided by Willmott Dixon Maintenance Ltd (see para. (d) above). Job descriptions and operatives' hours were transcribed to computer files, concurrently with analyses of the language used by their originators. Emerging patterns of sentence structure, ambiguities and contradictions were used to test the applicability of a minimum vocabulary; from these experiments it became possible to refine the problems and advance theories for their resolution.

Meeting the second objective involved investigation into the two computer languages, QBASIC and Prolog, chosen with the exigencies of this present research in mind, whilst technically acceptable in processing terms. Program modules were written that enabled output to be presented on a monitor screen. This approach allowed testing to be limited to single functions and immediately showed faults in logic or expectations. Programming did not extend to developing an entire system, for reasons stated above.

Practical considerations formed the background throughout this enquiry, and theoretical solutions were assessed against their general application.

3.

Language of job descriptions.

The grammar of our everyday language, whether written or spoken (and there are differences between the two modes) will not serve the functions of the automatic classification system now being sought. Therefore, it is logical to investigate the structures of sentences used in job descriptions, without becoming either too involved in theory, or enmeshed in the arcane discussions between psycholinguists and grammarians.

(a) Language system.

Linguists describe language as a system connecting meaning with sound and regard its complexity to be such that more than one level requires analysing. Transformational grammars, recursive and augmented transitional networks (RTNs and ATNs), Spreading Activation Networks, are but a few of the approaches in the development of modern theories of language structure and its connections with meaning. A more traditional structural breakdown recognises three main levels; phonology, syntax, and, semantics, divided between a grammar and a lexicon.

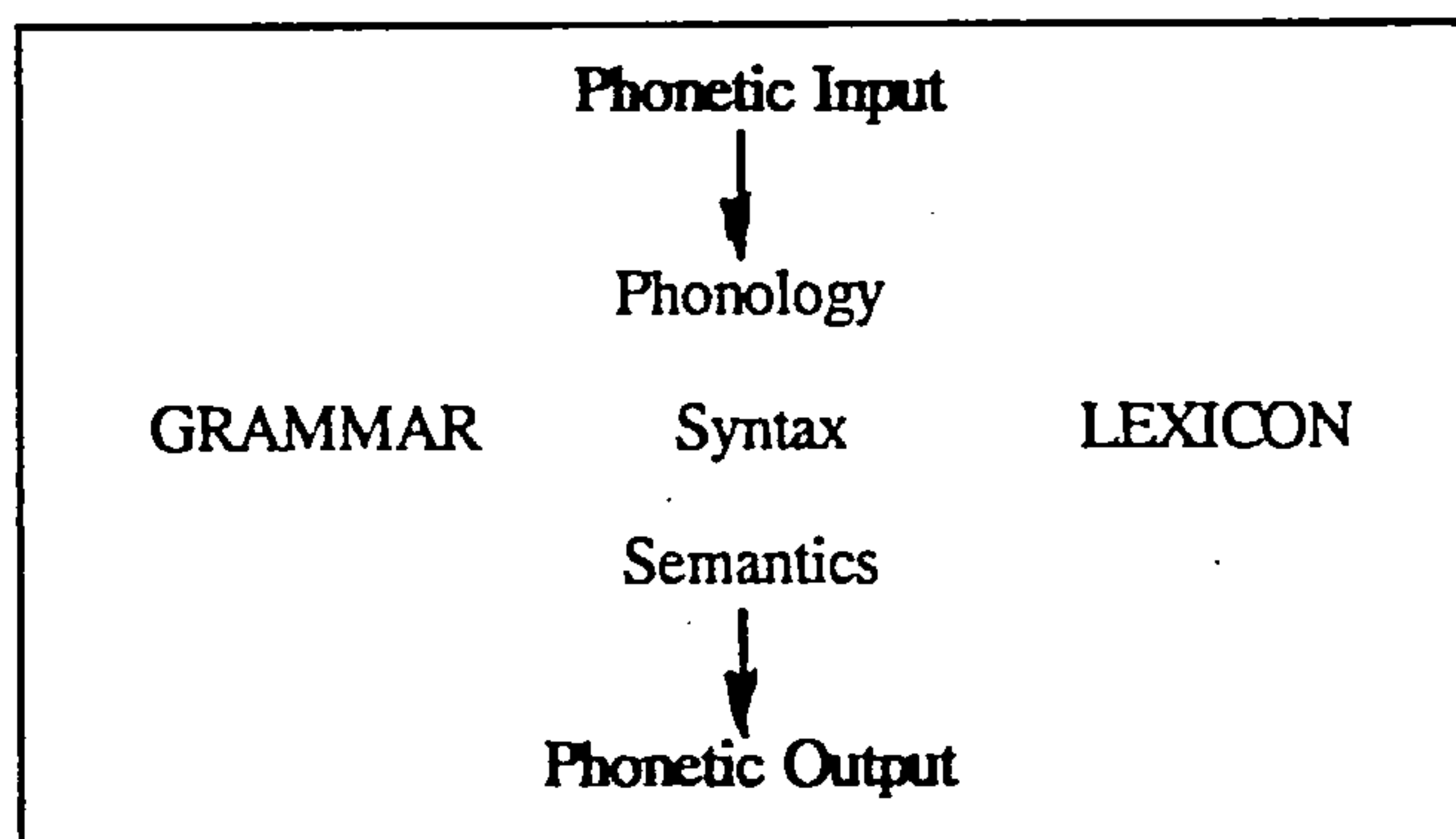


Figure 3.1

(After Leech, 1974)

This is a gross over simplification of the subject, but it serves a useful purpose in showing that continuum between 'grammar' and 'lexicon', and how it passes through the interchanging mechanisms of semantics, syntax and phonology.

Miller (1967) postulated a general hierarchy of the processes involved in communicating a sentence, and freely adapted to written work, becomes:

**Reading** A visual stimulus. Responses in terms of the legibility of the symbols, number, length of sentence, position on page.

**Matching** Recognition of the patterns of symbols in terms of the grammatical skills of the reader (or writer).

**Acceptance** As a complete sentence with a recognisable structure

**Interpretation** As a sentence meaningful in terms of its semantics, and, the collective and combinatorial meanings of the words it contains

**Understanding** Goes beyond interpretation as the information may not be written into the sentence, but be understood from the context, e.g., if part of a safety inspection report reads 'The ladder has two rungs missing' one would understand that the ladder was a potential danger if not in use, and should be replaced if it was being used. Adding the words, 'near the top', would provide the further information that the ladder was upright and if in use, the degree of danger would be increased

**Belief** Reader may believe a sentence is valid in terms of its relevance to his own conduct and knowledge. Thus the analysis of the sentence 'The plastering was nailed to the brickwork', would move through the hierarchy, level by level and then be understood, but not believed by anyone having knowledge of the skills and practices of plasterwork.

Rules of sentence generation are incorporated into a language once only, but their application is unlimited, which defines the starting point of this present research.

Sentences can be read by a machine, their words and structures matched with a lexicon and correlated to acceptable models, but can they be *interpreted*? The first step taken in seeking an answer to that question, was to establish the type of sentence-structures arising in the specialised language of building repairs.

Phonetics cannot be dispensed with entirely in written material, because they cause interference in word-matching, for example, when homonyms are confused. Consider *fowl* for *foul* in the sentence:

'To attend site and move position of door closer and refix so as not to fowl door.'

The sentence is acceptable, and may be *interpreted* and *understood* in three ways: *belief* only in the sense of a door being obstructed and not being made unclean by the door closer, or indeed having anything to do with poultry

(b) Structures of sentences in job descriptions.

Syntactical structures of sentences are shown as tree diagrams in Fig. 3.2 to Fig. 3.11, using a simple hierarchical convention composed of nodes, denoted in descending order. S for the sentence and clauses indicated as,  $S_1 S_2 \dots S_n$ . Constituent, or class phrases are categorised by a head word of the same word class which, as a general rule, is contained in the phrase. For example, in Fig. 3.2, the sentence, S, has four Verb-phrases (VP); *attend site, adjust regulator, allow door, close securely*. Without heads, each Verb-phrase will not be well-formed or grammatically correct, thus rewriting the sentence, with only its Verb-phrase complements, does not make sense, viz:

*to site and regulator on door closer to door to securely*

Similarly, for the Preposition-phrases (PP), the prepositions only exist because of the word-phrases they immediately dominate. Occasionally, a head word is not used, exemplified in Fig. 3.2 and 3.3, where the writers have omitted Determiners from Noun-phrases, *the* (Definite Article) and *a* (Indefinite Article). Where it is considered necessary, either for reasons of clarity or to properly construct a Class-phrase, the missing word is indicated by (e); it will be labelled if the originator's intention is beyond reasonable doubt.

The other word-phrases are, Noun-phrase (NP), Adverb-phrase (ADVP) and Adjective-phrase (AP).

Fabb (1992) raises an issue as to whether English can justifiably lay claim to the Verb-phrase "...with the verb, and its object, plus other things which follow the verb", (p.108)

but this argument has, after examination (due to Chomsky, 1988), been ignored because it introduces complications without advancing this present research.

Placed at the bottom level of any tree structure are nodes labelled with the component words of the sentence, and, immediately above them are their constituent categories, abbreviated:-

N = Noun, V = Verb, A = Adjective, ADV = Adverb, P = Preposition;

CONJ = Conjunction, M = Modifier, D = Determiner.

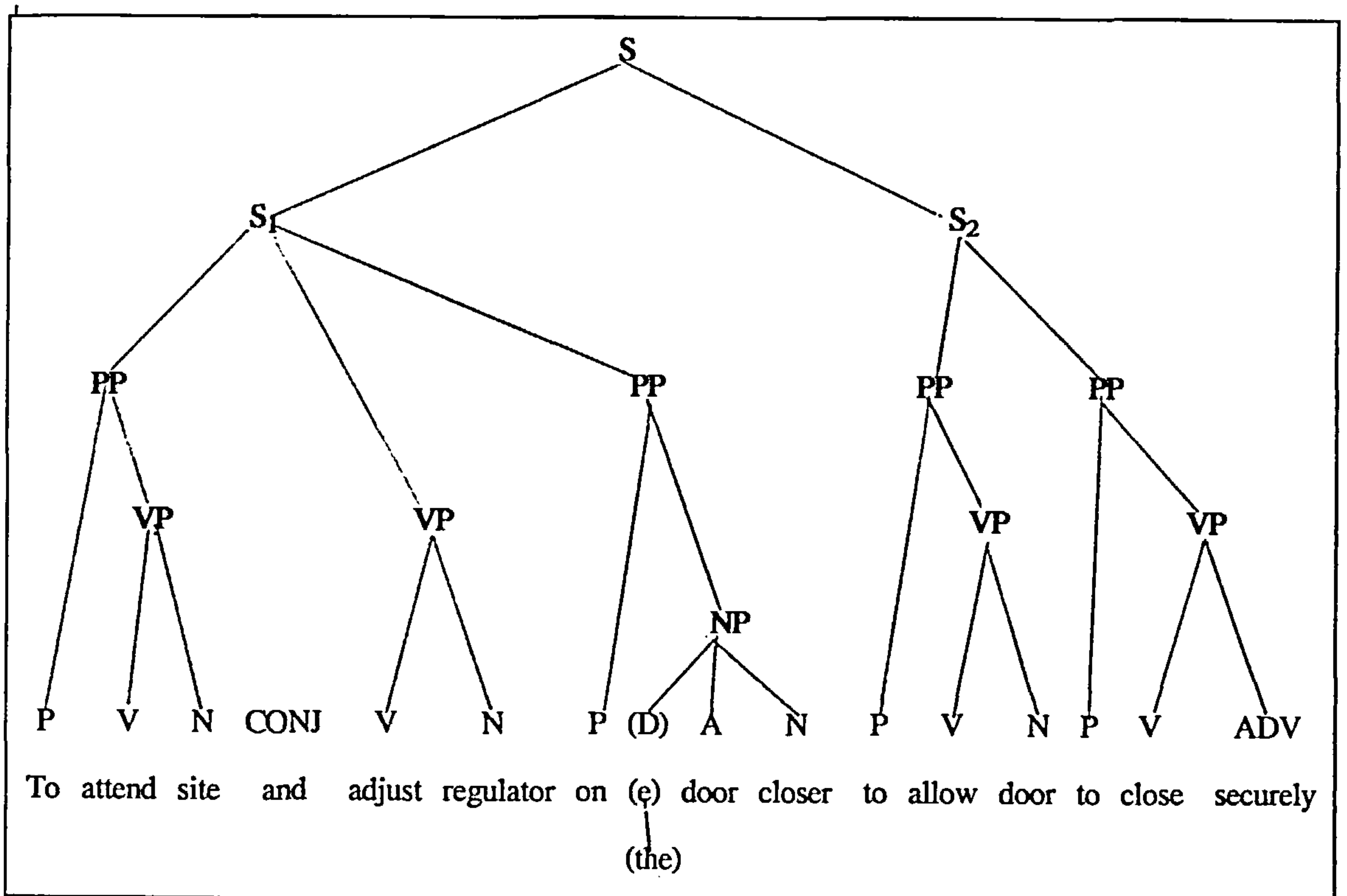


Figure 3.2

Returning to the syntax of the description under analysis, Clause S<sub>2</sub> if deleted in its entirety, along with the introductory 'place' Preposition-phrase, [To attend site], is still understandable as a maintenance task;

*and adjust regulator on (the) door closer*

The conjunction, *and*, is also unnecessary, so that in a reduced form it becomes,

adjust regulator on the door closer, S→ V+N+P+D+A+N

Changing the noun *regulator* to its verb form *regulate*, enables a simplification to,

regulate door closer, S→ V+A+N .....(1)

The sentence depicted in Fig. 3.3, was quoted earlier:

S<sub>1</sub> = To attend site and move position of door closer

S<sub>2</sub> = [a subordinate clause of S<sub>1</sub>] move position of door closer

S<sub>3</sub> = refix so as not to fowl door.

Note the phrase [and refix] conjoins S<sub>2</sub> to S<sub>3</sub> and could be omitted because S<sub>3</sub> or rather that part, [so as not to fowl door] is a constituent sequence, so that S then reads:

*To attend site and move position of door closer so as not to fowl door.*

Deleting the constituent sequence leaves:

*To attend site and move position of door closer and refix.*

and the reason for carrying out this task is not revealed

Dropping S<sub>3</sub>, as it is shown in the tree structure:

*To attend site and move position of door closer*

then *refixing* the [door closer] must be understood as the inevitable result of [move position]. Essentially, the same meaning is contained within the Verb-phrase;

*move position of door closer*

Thus, changing [move position] to *reposition* and deleting *of*, the sentence reduces to:

reposition door closer, S→ V+A+N .....(1.1)

However, an ambiguity is introduced. Without altering the word order, S→ V+N+A, the words *door* and *closer* have exchanged word classes, rendering the meaning,

"relocating a door nearer to some unspecified object or location; also, *door*, serves both

as adjective and noun in the original structure. (But are not the syntactical problems under discussion).

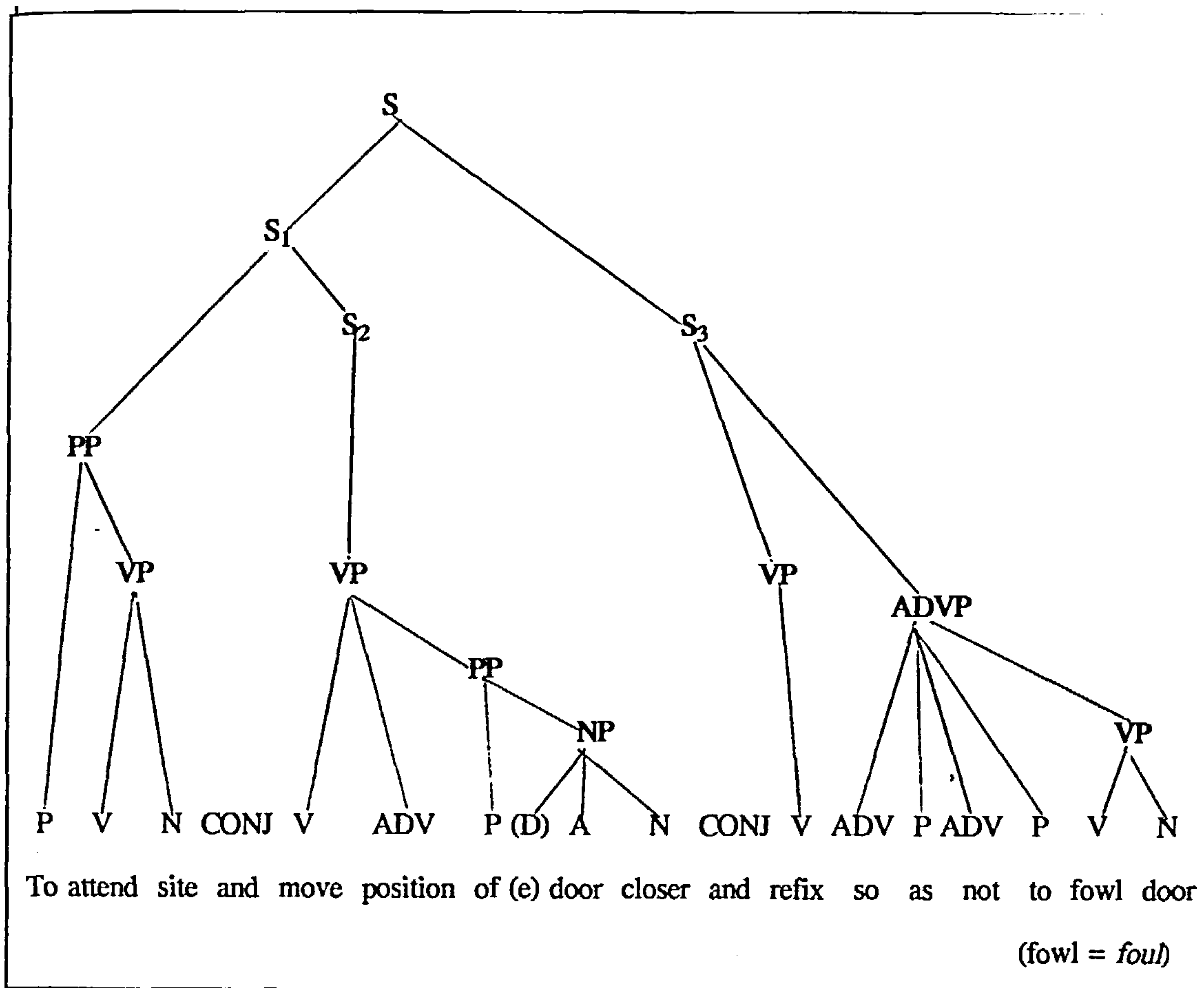


Figure 3.3

The next description (Fig. 3.4) is, superficially, a similar operation, once the tradesman is on site, having collected the new component:

*install after removing (the) old (door) closer.*

Missing head words may not cause problems in subsequent automatic sentence transformations provided their roles can be defined. The noun, *closer*, is modified by the adjective, *old*, which can only apply by referring to [door closer] in Clause,  $S_1$ . Deleting the idiomatic opening Preposition Phrase, together with both succeeding Verb Phrases, leaves:



\*door closer and install after removing old (e) closer

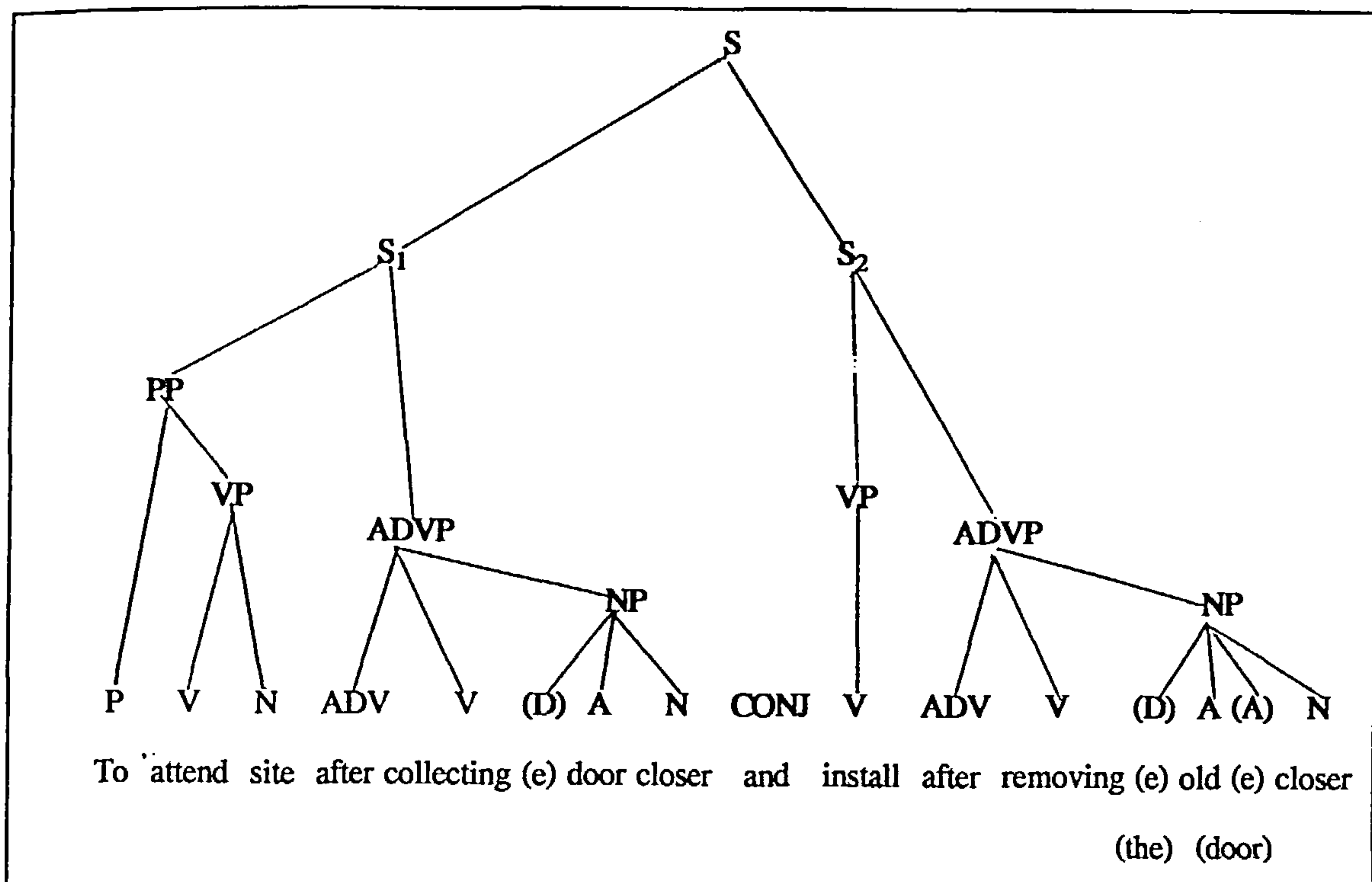


Figure 3.4

Word positions are an important feature of English grammar and this sentence is, therefore, badly formed, [door closer] being at fault. A grammatical sentence is obtained by re-arranging [door closer] to follow the verb *install*, although the head word, *and*, is not required for imparting meaning;

*and install door closer after removing old closer.*

A further reduction is possible by replacing *install* and *removing* with the synonymous, *replace*, (using *and* as a head word to start a sentence is still considered by some to be a solecism) ;

*and replace door closer after old closer, S → CONJ+V+A+N+V+A+N.*

A further simplification is, *replace door closer, S → V+A+N* .....(1.2)

or the slightly less ambiguous;

replace old door closer, S→V+M+A+N where M denotes, Modifier.

Alternatively, without a modifier, S→V+A+A+N .....(2)

The fourth tree structure diagram (Fig.3.5) displays yet another sentence describing the replacement of a door closer. Clearly, the first Clause, S<sub>1</sub> can be disregarded without further discussion, as it will be on future instances;

*remove bar entrance door closer, fit new (door)(closer) and check operation*

Some adjectives are superfluous, (*bar, entrance*) which only serve to identify one particular door from, possibly, several alternatives, the empty head word (*door*) in S<sub>3</sub> but not *new*.

*remove door closer, fit (closer) and check operation*

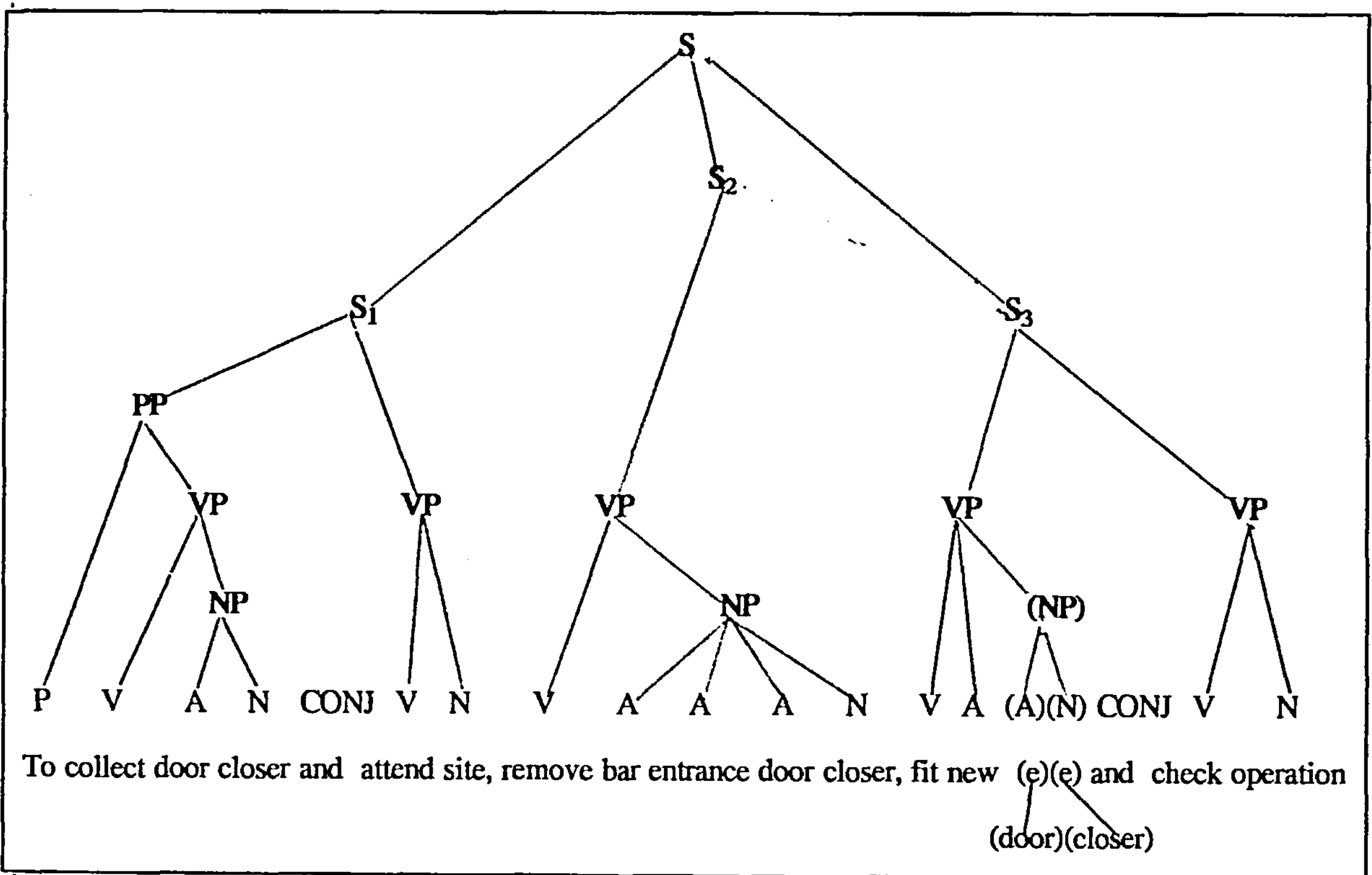


Figure 3.5

This sentence is, like a previous case (Fig. 3.3) somewhat ambiguous, again arising from the same word-class duality, *door closer*, A+N or N+A Redefining the comma as *and*;

*remove door closer and fit new and check operation*

"Checking the operation" after fitting a new component, particularly a hydraulic door closer, is inseparable from the action, *fitting*;

*remove door closer and fit new, S→V+A+N+CONJ+V+A+(N) .....(3)*

Then, as with the sentence in Fig. 3.4, the reduced formation becomes;

*replacing door closer S→V+A+N .....(1.3)*

Or,

*replacing door closer with new S→V+A+N+P+A .....(4)*

After making several assumptions about semantics, the resulting sentences describing a variety of tasks to do with door closers, are transformed by reductions into;

*S→V+A+N .....(1 to 1.3)*

Quite a different operation is shown in Fig. 3.6. The original job record carried additional information about the cause of the blockage and has been omitted, because it does not add anything to a classification or description of the activity.

The structure of the sentence is a cumulative sequence of Clauses;

$S_1 = \textit{rod all manholes}$

$S_2 = \textit{on drain run}$

$S_3 = \textit{to clear blocked WC.}$

S, could be constructed in either of these ways;

$S_{(A)} = S_1 + S_3 \textit{ rod all manholes to clear blocked WC}$

$S_{(B)} = S_1 + S_2 \textit{ rod all manholes on drain run}$

Since the number of manholes is unknown, the Determiner can be omitted. Stripped of irrelevant lexical items, the sentence reads;

*rod manholes S→V+N .....(5)*

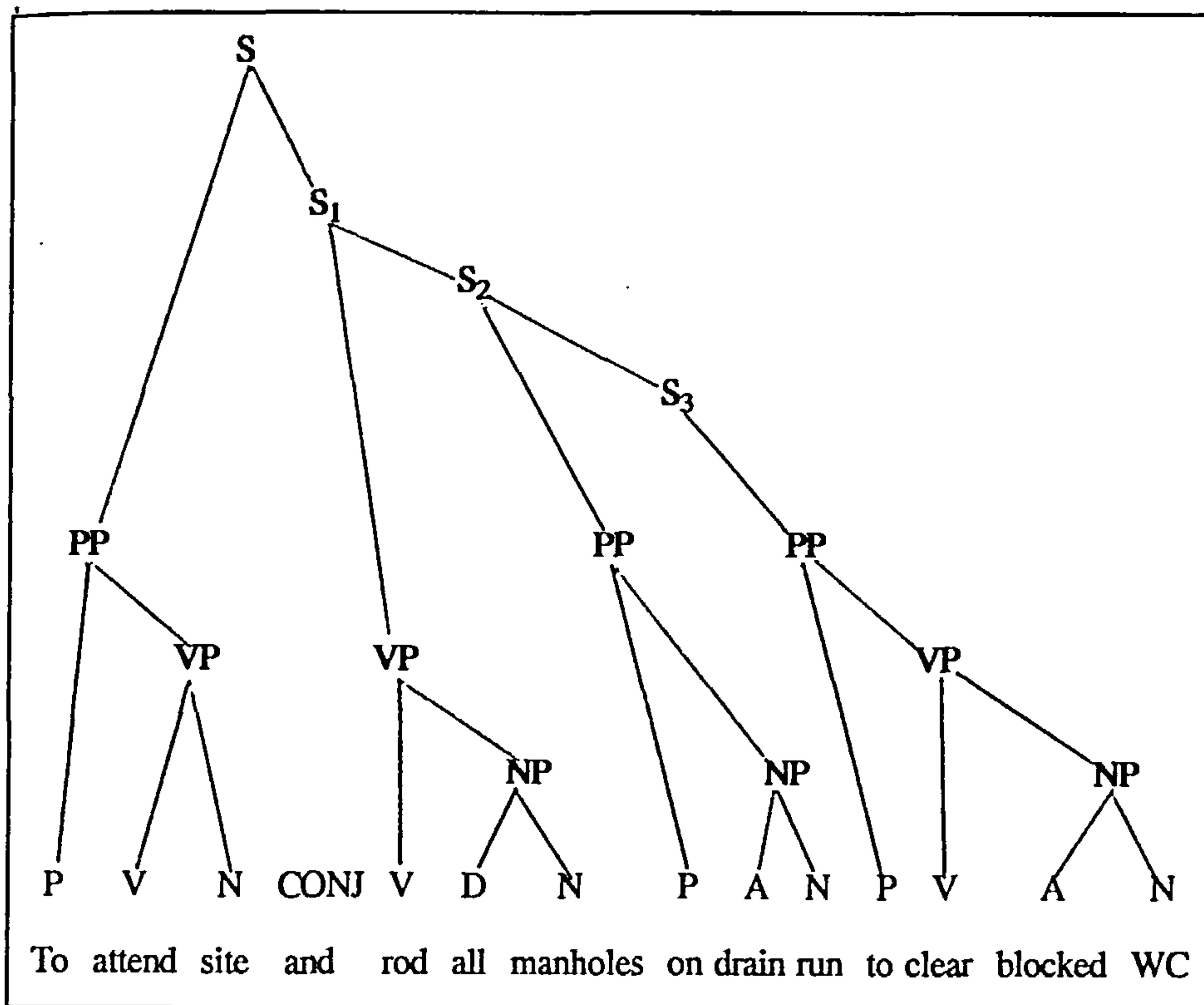


Figure 3.6

Yet another semantic duality is revealed: [rod] V in this context means the action of removing an obstruction from, or cleaning, a drain pipe. As a noun [rod]N can mean a metal or timber strip of uniform section, often part of a component or mechanical device.

Fig. 3.7 shows a construction of a sentence that might be describing two, separate tasks. Replacing a cylinder night latch, and, refixing a finger pull. Alternatively, and the more likely explanation, is for the finger pull to act as a face plate for the latch, being re-used with the new [Yale lock]. Syntactically, the sentence for reduction is;

$$S \rightarrow S_2 + S_3 = \textit{fit new Yale lock and refit (e) finger pull}$$

The Determiner is undefined, *a* or *the* are both feasible insertions.

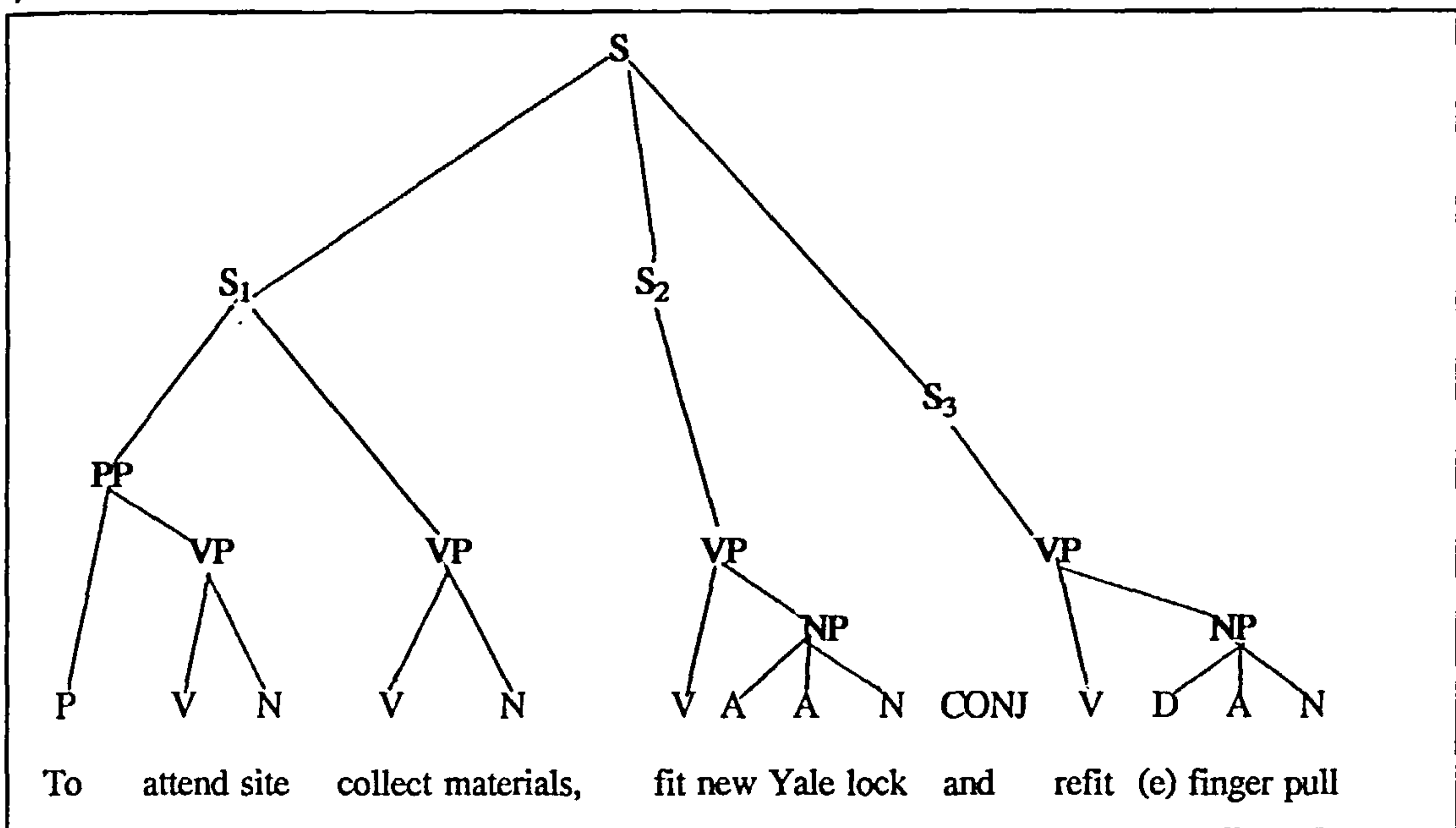


Figure 3.7

The adjective [new] is a *quality* and an unnecessary statement for the repair, leaving;

*fit Yale lock and refit finger pull*  $S \rightarrow V+A+N+CONJ+V+A+N$  ..(3.1)

and is a concatenation of the reduced sentence form (1 to 1.3).

The sentence in Fig. 3.8, opens with the usual Preposition-phrase, but this time, extended with a qualification, [on premium time]. This also can be ignored, allowing analysis of;

$S = S_2 + S_3$  *hack out (glass) and reglaze window adjacent to door*

Location, in the sense of where the window was situated, or which one was reglazed, is an unnecessary item of information, so the Adjective-phrase is deleted;

*hack out (glass) and reglaze window*  $S \rightarrow V+ADV+(N)+CONJ+V+N$

Removal of old or broken glass prior is an essential prerequisite to reglazing, thus allowing the deletion of the Verb-phrase concerned. The reduced sentence becomes;

*reglaze window, S  $\rightarrow$  V+N* .....(5.1)

and is the same reduced structure as for Fig. 3.6, "rodding manholes".

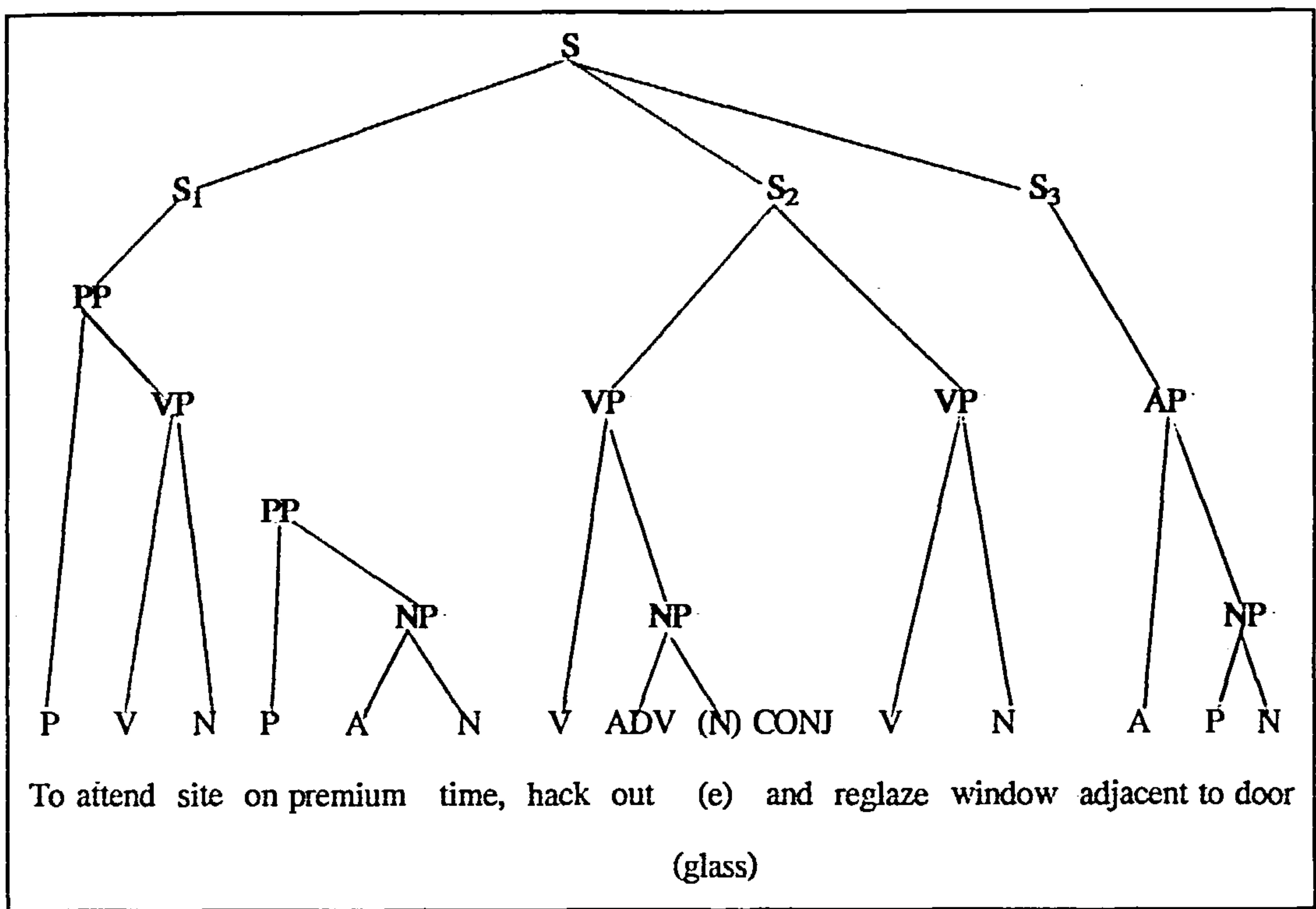


Figure 3.8

The next description (Fig. 3.9), is the most complex so far, with four Clauses and four empty lexical items. Following a similar sequence of deletions as taken previously, the sentence for reduction is:

*uplift loose quarry tiles, clean surfaces and rebed (quarry)(tiles), grout and polish off (tiles),*

$S \rightarrow V+A+N+CONJ+V+N+CONJ+V+(A)+(N)+CONJ+V+CONJ+V+ADV+(N)$  with every comma treated as a Conjunction. (The final missing word might be [grout]N, conforming with common practice when wall and floor tiling). Starting the reduction by deleting conjunctions, whether they are punctuations or words, together with empty lexical items, gives;

*uplift loose quarry tiles clean surfaces rebed grout polish off*

$S \rightarrow V+A+A+N+V+N+V+V+V+ADV.$

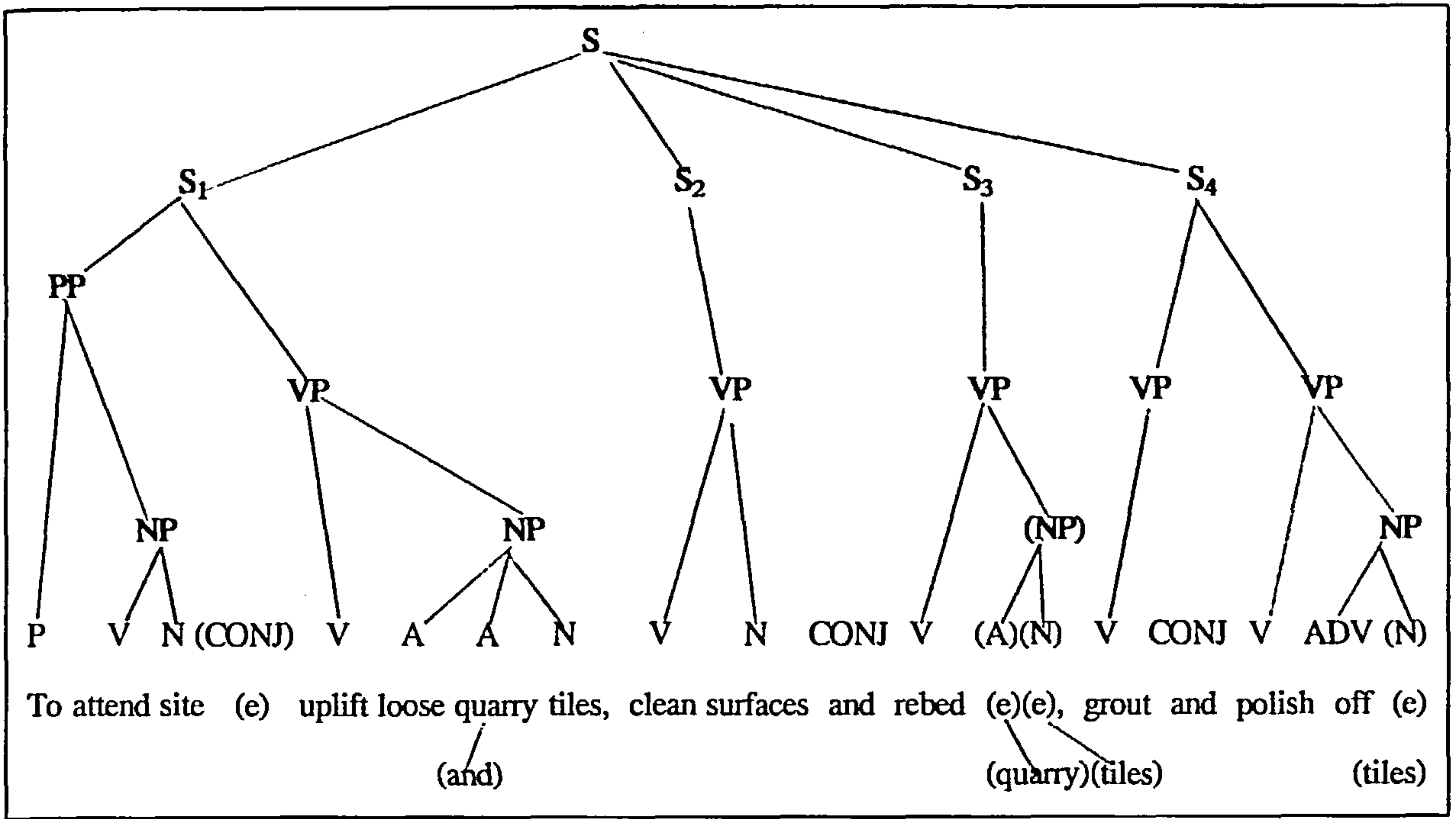


Figure 3.9

Clause S, provides the context for what follows, with the first adjective, [loose], deleted as an unnecessary qualifier. Clauses S<sub>2</sub>, S<sub>3</sub> and S<sub>4</sub> whilst constituents of the sentence S, they are dominated by S<sub>1</sub>, therefore any word-class other than Verb can be deleted from them. Either qualifier [loose]A, or [quarry]A, may be omitted from the first clause and the sentence becomes;

*uplift quarry tiles clean rebed grout polish or, uplift loose tiles clean rebed grout polish*

$$S \rightarrow V+A+N+V+V+V+V \dots\dots\dots(6)$$

and remains an understandable sentence, with its meaning imparted by the word sequence. A further reduction is possible by re-introducing the missing words [quarry] [tiles] to the Verb-phrase in S<sub>3</sub> and omitting the rest;

$$rebed quarry tiles, S \rightarrow V+A+N \dots\dots\dots(1.4)$$

Once again, the sequence S $\rightarrow$ V+A+N is the final transformation, but only after *inserting missing words*. Based upon deletions, the same sentence is attenuated to [uplift

quarry tiles], a very different task from [rebed (quarry)(tiles)].

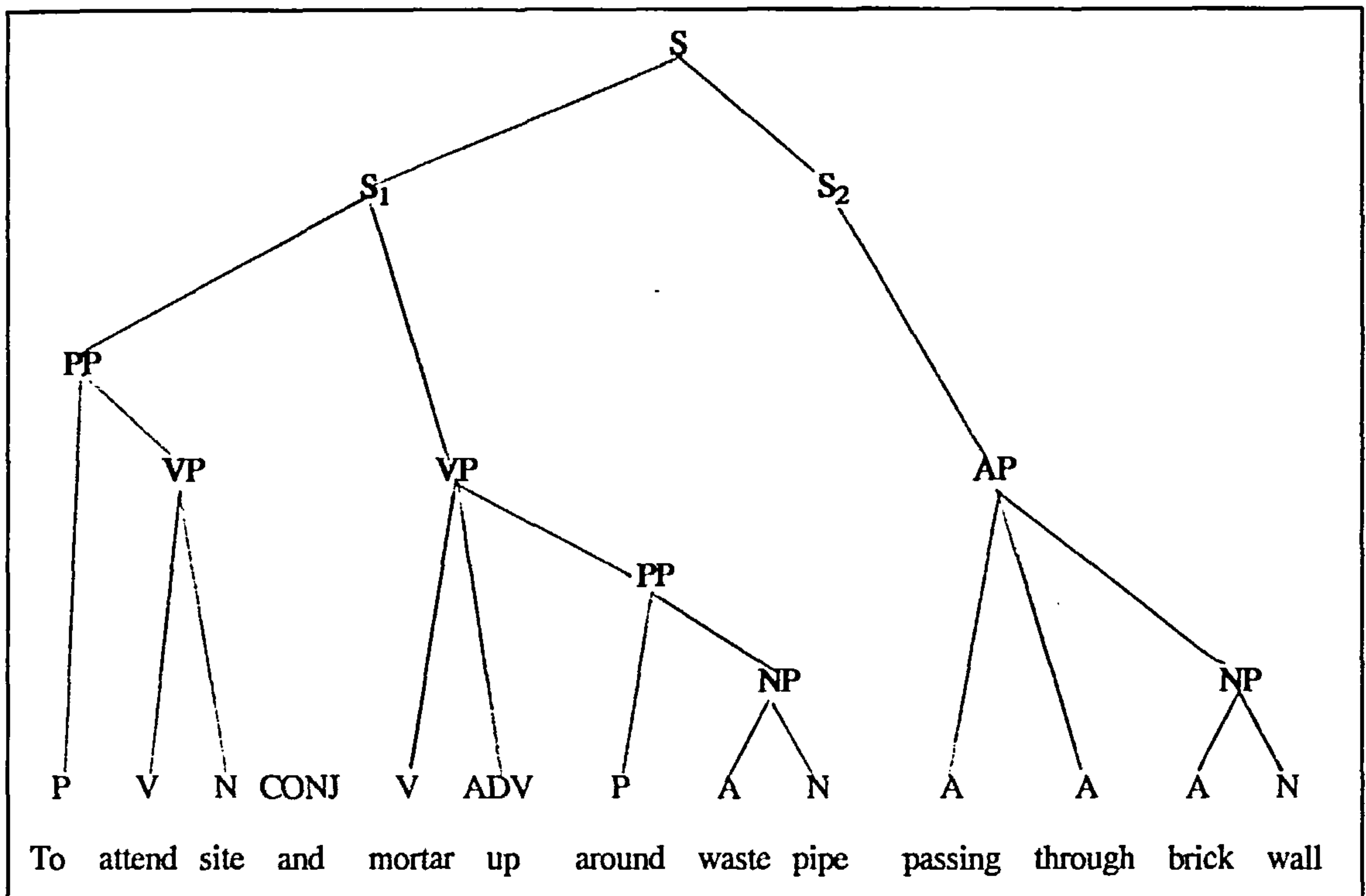


Figure 3.10

The penultimate sentence of the group under analysis (Fig. 3.10), records a simple repair, 'pointing brickwork around waste pipe', or, 'make good pointing around waste pipe', and, 'make good mortar jointing around pipe' etc. There are many other descriptions, equally adequate for defining a repair to a masonry wall where a small pipe protrudes from it. The term *point*, (or *pointing*) either with *brickwork* or *mortar* accurately describes a repair to an uncovered masonry wall surface. This definition is endorsed by *waste pipe*, which is any pipe discharging liquid horizontally through a wall, normally to the exterior of a building. Had the pipe been described with the adjective *vent*, some doubt would be aroused as to its orientation: 'vent' in the word-class Noun, removes any indication of size or cross-sectional shape.



The sentence is transformed simply by deletions to leave S<sub>1</sub>;

*mortar up around waste pipe, S→V+ADV+P+A+N*

Prepositions have, so far, been obvious candidates for removal. Here the result is a slightly vague description;

*mortar around waste pipe, S→V+ADV+A+N* .....(7)

The original meaning can be inferred only by labelling [mortar] as a Verb. Used as a Noun, the sentence becomes the incomplete, but grammatical form;

*mortar around waste pipe, S→N+ADV+A+N* .....(8)

Thus, leaving the reader asking "What was done with this mortar"? Deleting the Adverb from (7) offers the opportunity of constructing;

*mortar waste pipe, S→V+A+N* .....(1.5)

and an unsatisfactory description, especially when compared to the previous lexically filled version of (1).

The final description (Fig. 3.11) is associated with electrical work. Any term not directly related to the task can be removed,

*disconnect electricity supply take down sign remove*

*S→V+A+N+V+ADV+N+V*

The writer has not declared the type of sign, merely stating the disconnection of electricity supply as a prerequisite for its removal. There are several possibilities, two being; a neon sign, a sign illuminated by separate lamps, then either might be fixed to a wall or suspended from a bracket. Some kind of mechanically powered sign is another possibility. [Electricity] is necessary for a classification, but not for describing the work involved in the principal task;

*take down sign, S→V+ADV+N* .....(9)

The Verb-phrase [take down], implies a need to reach up and lower some object, or

component, rather than moving it. Syntactically, it is a well-formed sentence. A full transformation, with changes to both word-order and word-class, is;

*take down electrical sign, S→V+ADV+A+N* .....(7.1)

retaining it's definition as an illuminated sign and consequently, the necessity of making it safe before starting work.

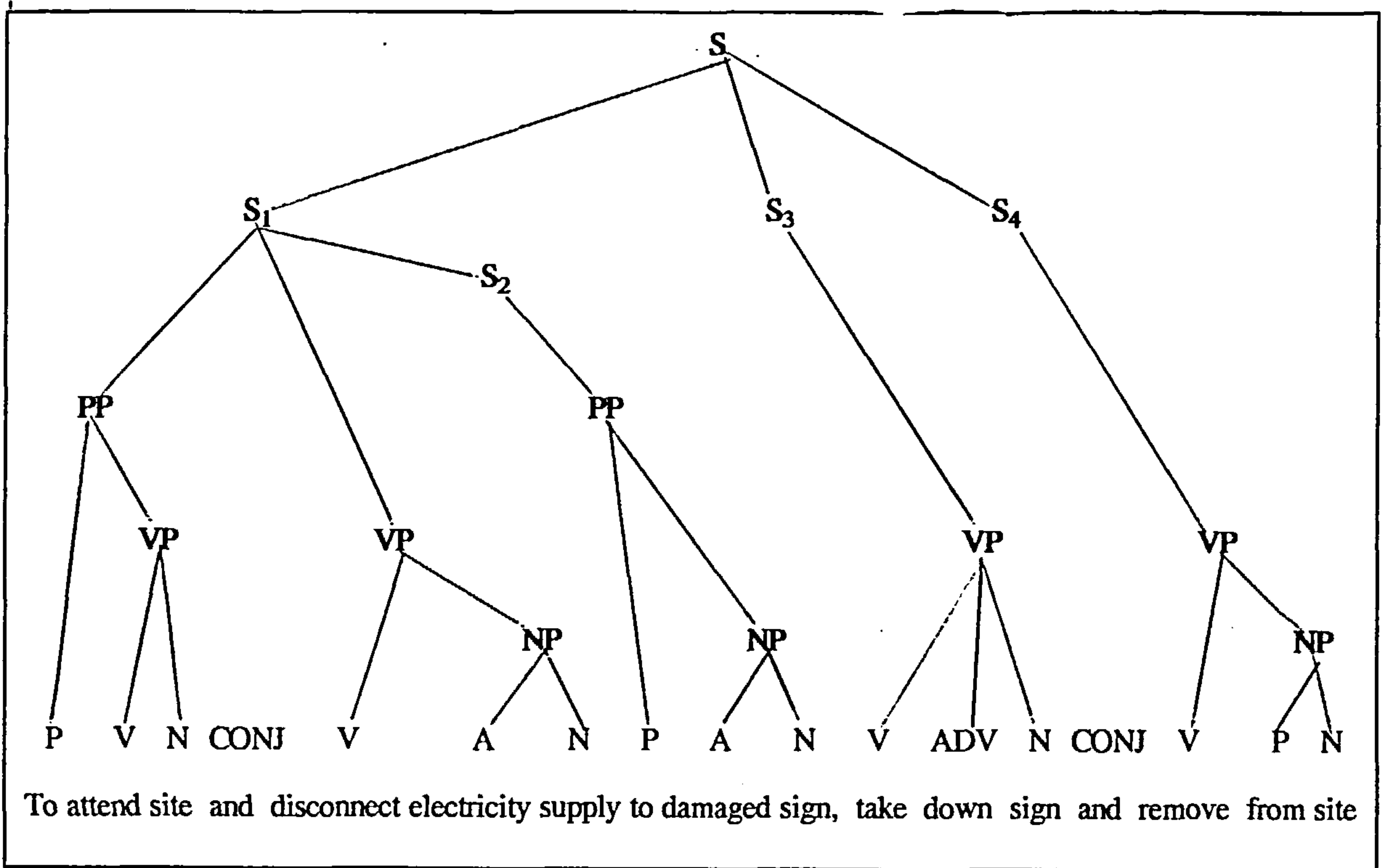


Figure 3.11

Attempting to force the sentence into construction (1), which has proved to be a commonly occurring outcome in this limited sample, involves exchanging [take down] for *removing* or *remove*, neither of which expresses the 'lowering' aspect of the original record;

*remove electrical sign, S→V+A+N* .....(1.6)

As a classifying term, it does offer some possibility that further examination might lead to it being applied in other sentence reductions, but there remains the difficulty of

automatically generating the necessary transformations.

(c) Narrative descriptions.

All the descriptions analysed so far were single sentences, written in an abbreviated form, and just sufficient to identify, for the client, the location and extent of the work being invoiced. Sometimes, records were written in a fuller, more discursive style, especially by workmen who carried out repairs involving a diagnosis of a fault, before deciding upon the best method to adopt in making a satisfactory repair. Here are four typical examples, taken verbatim from Day Work sheets and Confirmation Notes. (See Appendix D for examples) Note that the task sequence is given, with reasons, opinions and outcomes:

Attend site and find out which B/valve was overflowing. The float had broken off and the water was overflowing full out. Go to merchants and buy float. Return and drain of water from heating expansion tank, then remove b/valve to fit new float and replace b/valve.

Indicator light on - there showing solenoid had been engaged on stand by - alarm had been switched off from previous fault- (fault as in failure on float not electrical). Tried solenoid manually alarm & light on/off OK, but when in tank, float sometimes jams therefor alarm & light stay on whereas you can isolate alarm but light stays on - tried solenoid float in various positions in tank not very sucessfull. I therefor state main pump OK but standby (solenoid float) is only good for one warning - unless operated manually. System has not been tested fully when installed - would suggest if not bigger tank secure solenoid float in corner. Again main pump working - electrically OK.

Repair ball valve, find leak in soil pipe on pan connector. Remove pan, replace pan connector, refit pan for test. Found second leak on close couple, attempt to repair leak. No good. Too late to go to merchants, I will have to return. Go to merchants to get parts, return to site. Take out cistern, replace close couple washer, refit ball valve was jamming, so replace with new one, which had to be ajusted also. Test still leaking. Remove cistern again, ajust and refit, test OK.

Removed door from concealed overhead transome fixes closer & bottom pivot shoe, bottom pivot was worn/broken, tried to obtain pivot, pivot was old Briton 1100, which is now obselete. Briton pivot was centralised on metal diecast plate. Had to purchase new pivot along with new shoe as new pivots are offset from plate. Drilled new plate to refix into existing threaded metal base. Sawed 0.25" off new pivot. Redrilled shoe and reset pivot & shoe to suit. Adjusted both speed & latch adjusters as nec. Job completed

The first repair is the renewal of a float, entailing the removal of the ball valve [B/valve] [b/valve] and is described in the last sentence:

*Return and drain of water from heating expansion tank, then remove b/valve to fit new float and replace b/valve.*

$S \rightarrow V + \text{CONJ} + V + P + N + P + A + A + N + (\text{COM}) + \text{ADV} + V + N + P +$   
 $V + A + N + \text{CONJ} + V + N$

where (COM) = comma.

Reducing the sentence by extracting Verb-phrases, constructs the following:

*Return; drain of; remove b/valve; fit new float; replace b/valve.*

$S \rightarrow V; V + P; V + N; V + A + N; V + N$

Semi-colons are used to indicate the limits of phrases. A further reduction leaves a coherent abstract:

*fit new float and replace b/valve*  $S \rightarrow V + A + N + \text{CONJ} + V + N$

The sentence was selected intuitively, in the belief that to a knowledgeable reader, it conveys the essential task. Although the resultant form is syntactically correct, the lexical insertions provide relevance, making semantics the defining factor of a classification. Examining the 'disregarded' sentences, shows:

*Attend site and find out which B/valve was overflowing*

$S \rightarrow V + N + \text{CONJ} + V + \text{ADV} + D + N + V + V$

*The float had broken off and the water was overflowing full out.*

$S \rightarrow D + N + V + V + P + \text{CONJ} + D + N + V + V + \text{ADV} + \text{ADV}$

*Go to merchants and buy float.*  $S \rightarrow V + P + N + \text{CONJ} + V + N$

Using Verb-phrases from the previous analyses (Fig. 3.1 to 3.11), only  $S \rightarrow V + N \dots(5)$  matches in [Attend site] and [buy float]. The phrase "To attend site" is redundant, so "Attend site" qualifies for the same treatment. But turning a specific semantic exclusion into a general case, conflicts with inclusion by reason of syntax. Verb-phrases,

particularly when they are preceded by a conjunction, comma or as above, a determiner, may present a simple method of identifying those clauses worthy of a deeper analysis.

Accordingly, the three discarded sentences result in this concatenation of Verb-phrases:

*find out; B/valve was overflowing; float had broken off; water was overflowing  
full out; go to merchants; buy float.*

Only the phrase [find out] could reasonably be classified as a task, in the sense of it being a diagnosis of a fault for repair.

The next description defies the basic analytical processes followed so far. There is no focal repair, only a series of actions eliminating probable faults, leaving the positive information that the pump is electrically sound, as described in the extract:

*"I therefore state main pump OK but standby (solenoid float) is only good for one warning-unless operated manually."*

And reiterated in the last sentence, "Again main pump working-electrically OK". In passing, it is noted that 'OK'. 'Okay', 'o.k' or even 'oke' usually imply some form of test has been completed with a satisfactory outcome and may provide another phrase marker, similar to precedent conjunctions and punctuations.

The third job is a catalogue of tasks, linked together by explanations. Taking the sentences dealing with what can be identified, from the earlier analysis, as significant actions;

*Repair ball valve, find leak in soil pipe on pan connector.*

*S->V+A+N+(COM)+V+N+P+A+N+P+A+N*

*Remove pan, replace pan connector, refit pan for test.*

*S->V+N+(COM)+V+A+N+(COM)V+N+P+V*

*Found second leak on close couple, attempt to repair leak.*

S→V+A+N+P+A+A+(COM)+V+P+V+N

*Take out cistern, replace close couple washer, refit ball valve was jamming, so replace with new one which had to be adjusted also.*

S→V+A+N+(COM)+V+A+A+N+(COM)+V+A+N+V+V+(COM)+P+V+P+A+N  
+D+V+P+V+V+ADV

*Remove cistern again, adjust and refit, test OK.*

S→V+N+ADV+(COM)+V+CONJ+V+(COM)+V+ADJ

This selection of sentences, of course, begs the question, "What is a significant action?".

One response is to say, "Any action secondary to the objective task.", which by inspection gives:

*No good. S→D+A*

*Too late to go to the merchants, I will have to return.*

S→ADV+ADV+P+V+P+N+(COM)+N+M+V+P+V

*Go to merchants to get parts, return to site.*

S→V+P+N+P+V+N+(COM)+V+P+N

Syntactically, there is little to place these sentences in a different category to one, arbitrarily, called 'significant actions', (or 'relevant'), except for an absence of either Verb-phrase sequence, V+A+N and V+A+A+N. Another Verb-phrase is present, together with the Preposition-phrase, P+N [to merchant], [to site]. Then from this single observation, a sentence can be said to be without significance for a classification, if it does not contain either Verb-phrase structure V+A+N; V+A+A+N, but does include the Preposition-phrase, P+N. Formally expressed in program statements:

{English - the Sentence is significant if sig\_sentence{Sentence, yes}}  
significant\_sentence(Sentence):-

sig\_sentence(Sentence,X)!.}

:{English – the Sentence is not significant if it includes a PP and does not include a VP}

sig\_sentence(Sentence,no):-

includes\_PP(Sentence),  
not(includes\_VP(Sentence)).

Using only Verb-phrases and replacing prepositions and commas with semi-colons, the full sequence reads:

*Repair ball valve; find leak; Remove pan; replace pan connector; refit pan; test. Found second leak; attempt; repair leak. Take out cistern; replace close couple washer; refit ball valve was jamming; replace; had; be ajusted also. Remove cistern again; ajust; refit; test OK.*

Whilst there has been some reduction from the original, this compression obviously lacks the economy of most of the previous examples. Using the Verb-phrases 1, 2 and 9 to make a further selection in the order, (1), (1), (1), (9), (2), (1), gives:

*Repair ball valve; replace pan connector. Found second leak; Take out cistern; replace close couple washer; refit ball valve;*

Fortuitously, [refit ball valve] substitutes for the succeeding work, including [Remove cistern again], but misses the fact of the [refit ball valve] then being renewed. There is enough evidence to presume that an invoice writer might have rephrased the description, "To attend site. Renew defective ball valve. Find & repair leaks. Refit cistern. Test OK." However, generating a reasonably close match with an automatic transformation based on syntax alone, would be a considerable feat.

The last description of the quartet commences with a long sentence that contains a misspelt word, "fixes", which may mean "fixed" and a trade term, "Briton 1100", a piston operated door closer. There is also the ellipsis of *an*, [pivot was old Briton 1100],

therefore *an* is inserted and shown in parentheses. The eight sentences are constructed as follows:

*Removed door from concealed overhead transome fixes closer & bottom pivot shoe,*

*bottom pivot was worn/broken, tried to obtain pivot, pivot was (an) old Briton 1100, which is now obselete.*

S->V+N+P+V+A+N+V+N+CONJ+A+A+N+(CONJ)+A+N+V+A/A+(CONJ)+V+P+V+N+(CONJ)+N+V+(D)+A+N+(CONJ)+D+V+ADV+A

*Briton pivot was centralised on metal diecast plate.*S->N+N+V+V+P+A+A+N

*Had to purchase new pivot along with new shoe as new pivots are offset from plate*

S->V+P+V+A+N+ADV+P+A+N+CONJ+A+N+V+V+P+N

*Drilled new plate to refix into existing threaded metal base.*

S->V+A+N+P+V+P+A+A+N+N

*Sawed 0.25" off new pivot.* S->V+0.25"+P+A+N

*Redrilled shoe and reset pivot & shoe to suit*

S->V+N+CONJ+V+N+CONJ+N+P+V

*Adjusted both speed & latch adjusters as nec.*

S->V+D+A+CONJ+A+N+CONJ+A

*Job completed.* S->N+V

Since none of the sentences contain a Preposition-phrase, P+N, nor the 'significant action' Verb-phrase V+A+A+N, a selection will be made of those that contain the remaining phrase in the formula, V+A+N. The sequence is:

*Removed door from concealed transome fixes closer & bottom pivot shoe,  
bottom*



*pivot was worn/broken, tried to obtain pivot, pivot was old Briton 1100, which is now obsolete.*

*Had to purchase new pivot along with new shoe as new pivots are offset from plate.*

*Drilled new plate to refix into existing threaded metal base.*

The second sentence does not describe a significantly productive task, yet it conforms to the syntax rule by containing a V+A+N Verb-phrase, [purchase new pivot]. Once again, exclusion can only be made with the specific words, "purchase" and "new". Ignoring this semantic possibility and extending the selection with V+N Verb-phrases, reduces the sentences to:

*Removed door; concealed overhead transome; obtain pivot; was old Briton 1100.*

*Purchase new pivot;*

*Drilled new plate;*

*Redrilled shoe; reset pivot;*

Originally, the wording of the first sentence was opaque and it's reduction is ambiguous and misleading. Possibly, the writer intended to say, "Remove door from concealed overhead transome *fixing closer* (or *fixed closer*) and bottom pivot shoe", and certainly more likely than a hidden overhead transome. Only the Verb-phrase [was (an) old Briton 1100] indicates a repair associated with a door closer, whilst [obtain pivot] implies the acquisition was successful and not the reverse. What follows may not accurately describe the work actually carried out by its author, but is a pragmatic interpretation.

Remove door from concealed transome fixing and bottom pivot shoe. Renew worn bottom pivot, complete with shoe. Alter to suit Briton 1100 door closer, which is now obsolete. Regulate both speed and latch adjusters as necessary.

These narrative descriptions illustrate that although their individual sentences (apart from the second example) were amenable to analysis, interpretation depended upon referring semantic information from one sentence to another, something which is a unique feature of natural language. In practice, this is not likely to present an insurmountable obstacle to automatic translation, as invoiced descriptions are more usually paraphrased versions of the originals.

### Summary

This Chapter has shown, by manually analysing sentences, some strategies that might bring natural language of maintenance descriptions into the scope of computerised classification. Reduction in lexical choice is one approach that is advanced in the following Chapter.

4.

A minimum vocabulary.

The discussion has, so far, concentrated upon the style of English used in job descriptions, both that of workmen who carried out repairs and office staff when preparing accounts for their settlement. From the examples in the previous Chapter, there is little to choose between these two groups in terms of grammar and spelling, except for the latter who tend to adopt a more formal and terse phraseology which, to some extent helps in reducing sentences to limited sets of phrases. What is now proposed, is to carry succinctness further and examine the application of a minimum vocabulary in translating 'free' English into a basic sub-set.

(a) Restricting the choice of words.

Two problems found when manipulating the sentences in Figs. 3.2 to 3.11 (and the following, longer, records) were those of homonymy (Fig. 3.3) and synonymy (Fig. 3.4). That these are not isolated examples, nor the only forms of associations between objects and concepts, makes a method for mutual exclusiveness a necessity. Other researchers encounter similar difficulties, particularly because, '...the building sub-set is loose in construction and diverse in application' (Rougier and Lefley, 1994). They go on to cite Ruberg, Cornick and James (1989) who found in designing a system for the automatic diagnosis of building problems:

.a need for a universal taxonomy and pathology for buildings; experts in the building industry must come to agreement about the syntax of the problems they describe before a production system can be built to address the problem; (p. 399)

This present researcher did not have to consider, unlike all these authors, responses evoked from users consulting an expert system or a knowledge data-base. Here, there was no possibility of asking a writer to expand, or clarify, a job description. Neither could a syntax be imposed, at any stage, nor in the future, to ensure uniformity of

expression and prevent the use of idiomatic English. Therefore, in order to identify tasks for their *descriptions* to be classified, only relevant words have to be recognised and a minimum vocabulary offered the best mechanism for creating a gradus of mutually exclusive English words.

(b) A definition of a Minimum vocabulary.

Russell (1948) defined a minimum vocabulary as one containing no word capable of being defined in terms of other words in the vocabulary:

" Names are given to all the qualities of experiences, including qualities of visual space and remembered time. We also have to have words for experienced relations, such as the right and left in one visual field, and earlier-and-later in one specious present. We do not need names for space-time regions such as 'Socrates' or 'France' because every space-time region can be defined as a complex of qualities or a system of such complexes. 'Events', which have dates and cannot recur, are capable of being regarded as always complex; whatever we do not know how to analyse is capable of occurring repeated in various parts of space-time. When we pass outside our own experience, as we do in physics, we need no new words. Definitions of things not experienced must be denotations. Qualities and relations, if not experienced can only be known by means of descriptions in which all the constants denote things that are experienced. It follows that a minimum vocabulary for what we experience, is a minimum vocabulary for all our knowledge. That this must be the case is obvious from a consideration of the process of ostensive definition." (pp.282-3)

A minimum vocabulary to define, exclusively, maintenance work and similar short duration tasks, will not need words for 'experienced relations', or 'earlier and later', but it may require some qualities. A firm hypothesis could not be made on the evidence of a few translations, but there did not appear to be any serious disadvantages in a vocabulary that excluded,

proper names, for example of individuals, companies, trade-names;

prepositions and conjunctions;

events, for example, calendar, times, yesterday, today and the like;

subjective attributes, for example, colours, sound, relative dimensions..

(c) The Construction Industry Thesaurus.

The basis for such a minimum vocabulary, which at least, provided a building taxonomy, was a draft for the *Construction Industry Thesaurus* (Roberts et. al. 1970).

The Thesaurus (CIT) was arranged physically in two parts. First, an alphabetical Chapter serving three functions; the indication of synonyms with preferred terms; an index of relationships between terms; an indication of additional relationships to those set out in the second part, this being a systematic display of generic and whole/part relationships between terms. (Fig. 4.1). Every term used has a conceptual relation to the general body of information and these relations fall into eight categories or main facets:-

Time

Place

Properties and Measures

Agents of construction

Operations and Processes

Materials

Parts of construction works

Construction works

There is no need to analyse this structure, except to say that several facets failed to meet the criteria of a minimum vocabulary. Time, Place, Construction Works are obvious, but three seemed appropriate for maintenance, Operations and Processes, Materials and Parts of Construction Works. Some limitations of the CIT were found here: [Cleaning] and [Sewers], for example, being the only acceptable lexical insertions to represent a fairly common task - 'Unblocking drain'. There was no term for 'unstopping', 'unblocking' or 'unsealing', while [Sewers] was used for 'drains'. ('Used for' denoted a synonym in the CIT). Apart from plurality, 'sewers' implies something larger than 'drain', there is also

	(Shaped parts)	
	(Boards)	
J0791	Wood laminated boards	(Boards special to wood)
J0793	Laminboards	
J0795	Blockboards	
J0797	Coreboard wood laminates	
	(Joint covering)	(Boards by location)
J0799	Barge boards	
	(Windows)	
J0801	Window boards	
J0803	Featheredge boards	(Boards by form)
J0805	Plates	
J0807	Gusset plates *J1515	
J0809	Fish plates *	
J0811	Splice plates *	
J0821	Slabs	
J0823	Flagstones = Flags	
J0825	Tiles	
J0826	Slates	(Tiles special to material)
J0827	Flooring tiles = Paving tiles	(Tiles special to location)
J0829	Quarry tiles = Promenade tiles	
J0831	Roofing tiles	
J0833	Nibs = Cogs	(Parts of roofing tiles)
J0835	Laps	
J0837	Margins = Gauges	
	<u>Note</u> Divide like parts of roofs	(Roofing tiles by use)
J0841	Head nailed tiles	(Roofing tiles by nailing position)
J0843	Centre nailed tiles	"
J0845	Interlocking tiles	(Roofing tiles by joining method)
J0847	Lapping tiles	"
J0849	Plain tiles = Double lap tiles	(Lapping tiles by laying method)
J0851	Single lap tiles	"
J0853	Italian tiles	
J0855	Spanish tiles	
J0857	Pantiles	
J0859	Roman tiles	
J0861	Double Roman tiles	
J0863	Pooles tiles	
J0865	Walling tiles	

Figure 4.1 Extract from the Construction Industry Thesaurus (South Bank University)

the purpose of serving more than one property; the sewer being connected to drains, which serve properties on either side of a road. Even when the location of a blockage is specified, more difficulties can arise, as in [Cleaning] [Gullies], which does not properly classify 'unblocking gullies'. Further, all gullies could be classified by their function as [Interceptors], although the CIT in addition lists five traps by their form, at variance with the terms of Russell's definition.

(d) Consistency of word selection.

Having curtailed the CIT's reach, the principal difficulty in compiling a vocabulary, was framing rules to ensure a consistent selection of terms, both *from* the remaining displays and then for deciding what words needed to be *imported*. An initial approach, adopted by the present author in a previous work (Hague 1977), was to make several arbitrary decisions; plurals were made singular, except for products of uniform size; verbs were changed from the present tense to progressive (e.g. 'facing' to 'face', 'lifting' to 'lift'); some preferred terms were exchanged for the more relevant of the listed synonyms; 'drain' for 'sewer' mentioned above, 'repair' in preference to 'restore', and so on, without attempting to define 'relevant'. Rules for minimum vocabularies, like any other ordinances, can be construed in different ways, and within each construction there may be room for various degrees of conformity. For example, in Fig. 4.1, under [Roofing tiles] are tabled, (Parts of...); (... by use); (...by nailing position); (...by joining method), and, (Lapping tiles by laying method), it would be possible to use only two types, [Double lap tiles] and [Single lap tiles]. An equally good case could be made for defining the former as *plain*, with its synonym, *Rosemary*, a type of clay machine-made tile that frequently appeared in job descriptions. Alternatively, [Tiles] could be prefixed by [Roof], allowing other types of [Tiles] to be defined by a pertinent building element,



[Floor], [Wall] and [Ceiling], which then required decisions about appropriate synonyms. Concerns such as these have already been expressed in this present dissertation (and will be again); they do so because meanings of words are not only semantically important, the reasons for their importance shift with each context. Here, terms (words), were being selected for a vocabulary because, at a particular moment, they were perceived as instantiating things and actions involved with building maintenance. The question to answer was, 'Will these words be regarded as having the same meanings, to the same compiler after a period of time, or to others reading them as an automatically produced classification?' A philosophical analysis of 'meaning' lay outside the terms of reference for this present research, and certainly, it would have been unwise to risk becoming enmeshed in the fruitless endeavours described by Putnam (1975),

'...the dimension of language associated with the word "meaning" is, in spite of the usual spate of heroic if misguided attempts, as much in the dark as it ever was.' (p. 131).

(e) A strategy for selection of terms.

A general strategy was framed upon the proposition that a core of words must exist for denoting elements and sub-elements of buildings, which would, at the very least, enable operatives to be directed to a particular defect. There could be no equivocation or doubts about their intended meanings and usage, including misspelt words, such as, *rooves* for 'roofs', *eves* for 'eaves'. When descriptions tended to be specific, so doubts over the relevance *to this research* of the words used, increased with the degree of specificity. Proper names are already excluded under the rules of minimum vocabularies, (see para. (b) above) but the sentence "Repair door in Boys Block toilets on first floor", retains the potential to confuse if *Boys* and *first* are excluded. Further complications arise with the dual roles of *toilets* and *first*. In context, they identify a location within a building, separately, one is a component and the latter denotes a building element [suspended

floor]. Other examples of specificity can be difficult to translate when terms do not invoke the concept of a 'Place' or location. Typical cases are, [reveal] [borrowed light] [recessed/mortice/dead lock] [scrape] [strip] [polish] [roughen]. Points were, therefore, established where rules became necessary for the consistent selection of CIT terms (and imported ones). The following classes of words were identified:

parts of sub-elements and components (**Parts-of** terms)

types of elements, components, parts and materials (**Types-of** terms)

operations and processes..

Complexity increased with each level, indicated by the need for a separate Chapter to consider the rules for selecting verbs and verb-nouns, (See Chapter 5, 'A minimum vocabulary of operations and processes). It must be emphasised that none of the classes are clear-cut; some share words with common meanings, or terms are imbricated, where the meaning of one moves a short, but significant distance over the next.

(f) Part-of terms,~~X~~

A 'part' is defined here as being a piece of an entity, necessary to its *form* and *designated performance*. Thus, 'rafter' is a part of the element 'roof'; 'flaunching' is part of the sub-element 'chimney stack'; 'gutter outlet' is part of the system 'rainwater disposal'; 'float' is part of a ball valve, or [float operated device] in the CIT. The rule for selecting a **part-of** term is expressed:

a word denoting a part of an element, sub-element, component or system that can be replaced, or repaired, to restore the whole to meet its functional requirements, without first vitiating its form.

Under this rule, hierarchies of part-of terms are acceptable,

[washer, float, arm]–Part-of-> ball valve,

[washer, siphon, linkage]–Part-of-> siphon,

[ball-valve, siphon, overflow, cover, brackets]—Part-of-> cistern,

[pipe, cone-washer]—Part-of-> flush-pipe,

[cistern, pan, flush-pipe, seat]—Part-of-> water closet suite.

There are some parts, which, if detached from a body, or if a significant portion was removed, would lead an informed observer to infer that the whole entity required replacing. For example, 'reinforcement' (of concrete); parts of roofing tiles, (See Fig. 4.1) and bricks. Although roofing tiles are strictly parts-of the sub-element, 'roof covering', they are normally regarded as 'products of uniform size'. It was therefore, considered unnecessary to use terms denoting the integral parts of such entities, which are themselves renewed.

The exception rule is stated,

words or terms will not be selected for anything that may be denoted as being part of a product of uniform size, whether it is manufactured, or processed from a natural product.

However, care had to be exercised with exclusion rules for part-of terms; if this last rule had been widened to cover parts of all formed products, then 'reinforcement', 'aggregate' and 'cement' would not be admissible because they are indivisible parts-of the CIT term [Reinforced concrete], contradicting their primary role as materials. The words 'cement' and 'concrete' are tabled as materials under 'Synthetic inorganic complexes', then further classified, respectively, 'by property/composition' and, 'by parts'. They are also gerunds and their detailed consideration is deferred for the Chapter 5 'A minimum vocabulary of operations and processes.'

(g) Type-of terms.

Words for 'distinguishing a particular kind, group or class of...objects' (Shorter OED) by a quality, attribute, property or special feature were called **type-of terms**. Their selection seemed straightforward because a minimum vocabulary excludes words denoting qualities and relations that are not experienced. But what of words that in one sense are explainable by other words in the vocabulary, and are not in another case?

Some CIT terms have a class membership only discriminated by their context; [float] with [glass] has another meaning with [screed], [clear] with [drain] is not the same as when it is placed between [paint] and [varnish] or, [fix] and [glass]; [clear] [viewing spaces] has other (ambiguous) meanings from either 'unblocking' or 'transparency'.

Although very simple examples, they illustrate the problem created, amongst others, by polysemy. In addition, 'float' in 'float glass', 'clear' in 'fix clear glass' and 'clear' in 'paint clear varnish', when used as adjectives, express a feature of *glass* and *varnish*, as does 'perforated' *brick*, 'lightweight' *block* and 'gypsum' *plaster*.

Thus, a contradiction is revealed between the freedom to use any relevant English word in an association that confers a particular significance, and the proscriptive rules of minimum vocabularies. To some unquantifiable extent, a concept described by Putnam (1975) is of interest here. Where certain words have meanings known to one group of people and not to another, a phenomenon she called 'the division of linguistic labor' then arose. Her hypothesis was:

'Every linguistic community...possesses at least some terms whose associated "criteria" are known only to a subset of the speakers who acquire [sic] the terms, and whose use by the other speakers depends upon a structured co-operation between them and the speakers in the relevant subsets.' (p.146).

By adopting this proposition, it was assumed that anyone with a vocabulary of building terms and a knowledge of basic craft processes, would have the facility to discriminate

between classes, forms and the functions of building components, services and materials, at least to the extent of recognising when any term had been placed out of context.

(h) Type-of Building Elements.

Maintenance works with a duration of 16 hours or less, have limited scope for what can be achieved, so that elements of buildings could not become objects of descriptions, but remained the locii for repairs. Therefore, specific terms for differentiating types-of elements became redundant, including, for example, these CIT terms:

[Close couple roof] (UF Couple close roof)

[Collar beam roof]

[Lean-to roof]

[Mansard roof]

[Monopitch roof] (UF Single pitch roof)

This list covered terms for every roof form, even the participle, *pitched* and the adjective *flat*. Although both words re-appeared in the classification (Chapter 7) as generic divisions under Roofs, the convention that 'sloping' is 10<sup>0</sup> or more from the horizontal, (SMM7,1988; K20, Definition Rule D6) was assumed not to have been followed by the authors of job descriptions.

Applying this method for identifying redundant terms to other building elements and sub-elements, revealed the need to distinguish a specificity. Type-of terms for some walls define a purpose, as in the CIT term [Retaining wall], and also a position, 'sleeper' or 'fender' walls. (Dwarf walls supporting suspended ground floors.) Others also defined parts, such as 'diaphragm' and 'buttress', or more general words were used; 'single', 'double', 'curved', 'cylindrical'.

The rule for defining words in a set of **Type-of building elements** for exclusion was framed:

A term, or word, will not be selected if it is a name for a specificity unrelated to the principal purpose of the building element or sub-element it characterizes.

(j) Type-of Components.

Words for **Types-of components** were excluded if the prefixed term did not change one or more of the product's 'expressed purposes'. (CIB,1972). For example, *windows* are tabled in the CIT as [Eyebrow], [Bay] in various styles, [Dormer], [French], [Rose], [Combination] and there are others not listed, but none relate to the purposes of *windows*. An important distinction for classifying **Type-of windows**, is their *action* [Sliding], [Reversible], [Casement] (implying *hinged*). Similarly, for doors, there are CIT terms for [Revolving], [Folding], [Rolling], but not 'up-and-over'. Another illustration of where context caused word-classes to overlap was the word, *flush*. A *flush* door is a conventional type, usually defined by its fire resistance and facing materials, the hinged action is implied, being understood in the absence of any other; *flush* does not change the purpose of a door, any more than would *panelled* or *glazed*. The term [Flush] must be kept in the vocabulary for insertion as an operation, (verbs), and rejected as a **Type-of term** (adjectives); it cannot be a **Part-of term**, because *flush* does not belong to the word-class, 'noun'. Such anomalies and ambiguities became more frequent as the types moved away from building elements to the greater variety and complexity of materials, making their selection less clear-cut.

The rule for **Type-of component terms** was:

A term, or word, will not be selected when it describes a specific form of component unrelated to one or more of the expressed purposes of the product. Exceptions are made for terms defining an action (normally a directed motion).

Under this rule, [Close couple] as a **Type-of** water closet was rendered inadmissible, confirming its previous redundancy as a **Type-of** [Roof]. It also ensured the redundancy of adjectives such as *flush* and materials used to manufacture the components.

Uncertainty was further reduced, when a term from one level could modify one from another. Examples of this are; floor tiles, door closer, manhole wall, roof light, but there was no assurance of the CIT itself being complete, or that operating the rules presently under consideration would prevent the importation of relevant terms.

(k) Polysemy.

Logic and commonsense demanded that polysemic words had to be considered as a class applicable to all **Type-of** terms, particularly components and materials, whether or not they have been included by the CIT.'s compilers. The difficulty presented by polysemy for any lexicon or minimum vocabulary can be shown by the word *concrete*. It is a material and a noun in the sentence; *Place concrete*, an action and a verb in; *Concrete post holes*, a **Type-of** description and an adjective in; *Fix concrete posts*. Unless a differentiation is imposed, it would be possible to reduce this invented sentence: *To breaking out and concreting new 8ft. concrete fence posts. into; remove concrete; concrete 8ft. concrete fence posts.*

A general rule for limiting the word-classes for polysemic **Type-of** terms can be stated:-

'When any **Type-of** term is polysemic, then only the noun, verb and adverb senses are classifiable and not the adjectival.'

Having such an axiom raised the question as to what purposes are served by Type-of terms? If a word can be dispensed with, solely by reason of being polysemic, then what justifies the use of other adjectives? Deleting the adjective, *concrete*, from the immediate example, left a badly formed, but classifiable sentence, *remove concrete; concrete 8ft fence posts*. Regardless of deletions, the original persists as meaning, 'Remove concrete and fix new 8ft. fence posts set in concrete.' Other materials or different kinds of concrete posts have not been denoted, so implications of costs are irrelevant. However, *concrete* being a dense material, the posts have an intrinsically higher work-content, involving more effort (and cost) than would timber posts of similar dimensions, leading to a possible justification for Type-of terms beyond that of defining elements.

(1) Cost as an attribute of selection.

Cost, as an attribute of a job description was not considered in this research, but there were two aspects that were generally denoted by adjectives; the simple, directly comparable costs of materials of manufacture; and as exemplified by *concrete*, indirect costs of labour arising from variations in work-content due to the component's physical characteristics. Both are subjective judgments of a single question, "Would this term, if used as an attribute adjective or object complement, suggest an influence upon the labour cost?" If the answer was "No," then the term was inadmissible, and, *a priori* admissible for "Yes".

One example, discussed previously, concerned 'clear', which is widely understood to describe the quality of transparency common in; float glass, toughened float glass, clear laminated safety glass and Georgian wired polished glass. None of these descriptions, including the past participles 'toughened' and 'polished', denote significant factors beyond



differences in their costs as materials. Wired, or reinforced glass, does have a reputation for being difficult to handle, although in that respect, is hardly different from any other glass obtained ready to fix. Glass thickness and sheet dimensions, however, do affect the work-content, but should a distinction be made? And if so, where? Thicknesses of 2, 3, 4 and 5mm, with little to choose between them, are the lower end of a range that steps in larger increments ending at 12mm, but nothing to justify a separate term.

Semantically, dimensions are nothing more than tokens for different numbers rather than **Type-of** terms for glass. Operationally, work-content is dictated by the site conditions, together with the weights and sizes of individual panes; it is described by the preparation and glazing methods used, and, ultimately, expressed by the recorded job times.

Cost, therefore, is shown to be an unreliable indicator for selecting a term, or deciding between possible alternatives. Neither does it nullify the 'axiom of polysemy'. The earlier example of a [Close couple] [Water closet suite] was confirmed as, [Water closet suite]. Nevertheless, operatives often write details about the components that could be associated with work-content and it seemed worthwhile responding by allowing **Type-of** components to be identified by its material, if given in the description. Then baths that are supplied in various shapes and sizes, would be 'plastic' (for acrylic), 'porcelain enamel' or 'cast iron', and windows as 'softwood' or 'aluminium'.

(m) Type-of materials terms.

**Type-of material** terms were defined by their generic denotations if processed products of uniform size, (bricks, blocks, tiles), and by their expressed purpose if products of uniform section (pipes, gutters, architraves, channels). The remaining natural and processed products were then denoted by their generic name, (hardcore, stone, glass,

plaster, copper, lead, paint).

The rule for selecting a **Type-of material** term is:

the generic name of a material or expressed purpose of a product in their merchanted forms.

Whether or not materials would be prefixed by other **Type-of materials** terms has been considered elsewhere, but it can be assumed without defence at this point. Thus, examples of acceptable paired terms are, *plywood* sheets, *blockboard* shelf, *hardwood* sill, *cast iron* pipes.

(n) Summary and application of selection rules.

The question of polysemy was still unresolved. The redundancy of adjectives might be exchanged in the 'axiom of polysemy' for another word-class, and this is argued in the next Chapter, 'A minimum vocabulary for operations and processes.' Nothing better exemplifies the complexity of this process, than in words belonging to more than one class, for example, *loose* (A,V, ADV, N), *loosing* (V, ADV, N), *loosen*(V). The CIT. (prior to changing the verb-form herein) prefers [Unfixing] for [Loosing. (*sic*)], (see Fig 3.9), but *loosen*, or *loosening*, are now common parlance for an action bringing about that loss of firmness of attachment called *looseness*. Further, *unfixed*, a participial adjective, denotes *detached* as does *loose*, in the sense of, 'Not joined to anything else.' (Shorter OED.). But *loosen* (or *loosening*) does not mean *detach*, which is tabled in the CIT under the sub-heading, 'Separating by method' - [disconnecting = disjoining = uncoupling = detaching]; neither does it mean, *take apart*, a phrase associated semantically with [dismantling], a separate term denoting the breaking down into constituent parts, of a component, element or structure.

Summarising, the rules for selecting terms so far postulated are:

a word denoting a part of an element, sub-element, component or system that can be replaced, or repaired, to restore the whole to meet its functional requirements

a term, or word, will not be selected if it is a name for a specificity unrelated to the principal purpose of the building element or sub-element it characterizes.

a term, or word, will not be selected when it describes a specific form of component unrelated to one or more of the expressed purposes of the product. Exceptions are made for terms defining an action (normally a directed motion)

the generic name of a material or expressed purpose of a product in its merchanted form.

These rules were tested using the descriptions previously analysed (Figs. 3.2 to 3.11) as being a small, but representative selection. The sentences were first matched with terms listed in the CIT Sections: 'Operations and Processes'; 'Materials'; 'Parts of Construction Works', and then with the parsed, transformed or reduced sentences.

*To attend site and adjust regulator on door closer to allow door to close securely.*

[regulator USE automatic valve] [door] [closer] [door]. Similarly, the reduction,  $S \rightarrow V+A+N$  (*regulate door closer*) only matches [ ]+[door]+[closer] as there are no terms for *adjusting* or *regulating*.

*To attend site and move position of of door closer and refix so as not to foul door*

[moving] [position USE place] [door] [closer] [door] The closest match to *foul* is [foul gas USE crude gas.  $S \rightarrow V+A+N$  (*reposition door closer*) is matched [ ]+[door]+[closer]. (The CIT term [relocate] does not have any synonyms listed). Also, it is noted that the axiom of polysemy is followed because *closer* is a noun by its inclusion in the CIT Section, 'Parts of Construction Works'.

*To attend site after collecting door closer and install after removing old closer.*

[door] [closer] [remove UF clear] [closer]. S→V+A+A+N (*replace old door closer*) only matches [ ]+[ ]+[door]+[closer]

*To collect door closer and attend site, remove bar entrance door closer, fit new and check operation.*

[door] [closer] [remove] [door] [closer] [operation] S→V+A+N+CONJ+V+A (*remove door closer and fit new*) matches each word except the conjunction and the last adverb, [remove] [door] [closer] [fit] and the further reduction, S→V+A+N (*replacing door closer*) matches every term in, [replace] [door] [closer]. Again, the axiom would exclude the adjectival form of *closer*.

*To attend site and rod all manholes on drain run to clear blocked WC.*

[rod] [sewer] [remove UF clear] [block UF plug stop]. S→V+N (*rod manholes*) can be matched in [rod] only because this CIT term has the characteristic of 'Parts by Shape', and is not classified as an operation or process. There are no entries for *manhole(s)* or any of its synonyms, *access* or *inspection chambers*, the closest being [pits] in the Section 'Construction Works'.

*To attend site collect materials, fit new Yale lock and refit finger pull*

[fit] [ ] The CIT, as would be expected, eschews proper names, so a user must, in this instance, select the term [night latch] S→V+A+N+CONJ+V+A+N (*fit Yale lock and refit finger pull*) also only matches [fit].

*To attend site on premium time, hack out and reglaze window adjacent to door.*

[window] [door]. The CIT makes no provision for the process of *hacking out* although [hack USE roughen] is included; [glaze] appears with 'Materials for protection and decoration by use'. The reduction,  $S \rightarrow V+N$  (*reglaze window*) can only match with [window], raising the question of including in the CIT, several words prefixed 're' that describe operations or processes.

*To attend site uplift loose quarry tiles, clean surfaces and rebed, grout and polish off.*

[quarry tile] [clean] [face UF surface] [grout] [polish]. There were two forms of lexical insertions for the reduced sentence,  $S \rightarrow V+A+N+V+V+V+V$  (*uplift quarry tiles clean rebed grout polish*), and, (*uplift loose tiles clean rebed grout polish*). A match could be forced by using [lift] for *uplift* and [place UF lay locate position] for *bed* or *rebed*, then the sequences of terms are, [lift] [ ] [quarry tiles] [clean] [place] [grout] [polish], and, [lift] [loose USE unfix] [floor] [tiles] [clean] [place] [grout] [polish].

(Note the need for plurality in [tiles], referred to earlier).

*To attend site and mortar up around waste pipe passing through brick wall.*

[mortar] [waste pipe] [brick] [wall UF vertical dividing elements] The CIT also includes [brickwork] with sub-classifications of 'Brickwork by face characteristics', 'Brickwork by thickness' and 'Brickwork by bond characteristics', all of which can be discarded under the rules of a minimum vocabulary. The transformed sentences match [mortar] [waste pipe], but then in the form  $S \rightarrow N+A+N$ , not  $S \rightarrow V+A+N$ , because in the CIT [mortar] is a noun and *waste pipe* is one classification term and not two.

*To attend site and disconnect electricity supply to damaged sign, take down sign and remove from site.*

[disconnect] [electricity supply services] [sign] [sign] [remove]. The two reduced sentences, S→V+ADV+N (*Take down sign*) and S→V+ADV+A+N (*Take down electrical sign*), can only match [sign].

### Results.

These simple experiments lead to three resolutions.

First, as demonstrating support for the exclusions declared earlier in this Chapter at (b) A definition., of proper names, prepositions and conjunctions, events, and subjective attributes.

Second, how the demands of this present research are at variance with the objective of the CIT, stated as:

'...the development of an authority list of terms for the construction industry taking into account the need for different levels of application, and the requirements of other methods of information handling. (Roberts et. al., *ibid.*)

and exposed its antecedents as a vocabulary for indexing and information retrieval.

Third, that focusing upon the properties of words still seemed to hold the best promise of influencing the retention of meanings of sentences, expressed and implied, following their machine-translation into a limited sub-set of English.

## Summary.

Although words could be excised by applying the rules for selection, pragmatic decisions still had to be made, with at times contradictory results. The CIT, given its indexing bias, remained the principal source for compiling vocabularies, acting as a standard against which terms could be examined for relevance. It did not, by itself form any lexicon or vocabulary. The next Chapter develops the application of a minimum vocabulary to verbs and analyses their lexical and semantic roles.

5.

A minimum vocabulary of actions and processes.



This Chapter develops the proposition that difficulties of classing repair work by machine can be largely overcome by limiting the range of lexical choice. The analysis centres upon the role of verbs according to the object of a job description, which was, in part, examined in Chapters 3 and 4, together with further exploration of resulting ambiguities.

(a) Conflicts of meaning and recurring phrases.

The reduction of sentences made by applying the principles of minimum vocabularies, were a rehearsal of the problems envisaged by Nagao (1989) for machine translation;

'Deciding on the correct translations for verbs is an even more difficult task than that for nouns. It is known that many subtle factors can influence the choice of the appropriate verb, and such complex processing is still beyond the capacities of machine translation systems. In general, the selection of the verb is made on the basis of the meanings of the subject for the verb and its direct and indirect objects.' (p.110)

Fortunately, there are profound differences between Japanese and English (Nagao distinguished thirty-three cases in the analysis of Japanese verbs) but the relative simplicity of English still leaves room for linguistic puzzles. The gerund, or verb noun, is an example cited by Campbell (1982) in his critique of information theory; 'A sentence like "John's smoking bothers me"...appears to contain a possessive and a verb, yet in reality contains neither. The ambiguity is unresolved in spite of years of patient effort.' (p. 257)

However, verb-phrases tend to a consistent form, a fact observable in Chapter 3, where parsed sentences reduced to:

S→V+A+N                      ... (1)

S→V+A+A+N                    ... (2)

S→V+A+N+P+A                ... (4)

S→V+N ... (5)

S→V+A+N+V+V+V+V ... (6)

S→V+ADV+A+N ... (7)

S→V+ADV+N ... (9)

Most frequently used is the verb-phrase, S→V+A+N, which, in familiar language means 'An action or process carried out upon a type of object.' presaging a grammatical simplification of the classifying language. Rules for selecting terms have caused the redundancy of many adverbs and adjectives, increasing the probability of this pattern, and S→V+N occurring with a similar frequency in other, larger, samples.

Devising similar rules for selecting verbs, principally from the CIT, appeared to be replete with complications, beginning with a conflict between two interpretations of *processes*.

A note prefacing the schedule, Operations and Processes (See Appendix A), explained that they 'involve human action', and:

'...are actions which occur within materials..as a result of internal dynamism...usually without human involvement.'

This narrow view opposes that usually accepted in the building industry, but with a typical definition, implied by Pilcher (1992) in his description of process charts, which:

'...provide diagrammatic means for recording the sequence of activities in an existing method under study.' (p.209)

He reinforces the relevance of *process* in building work as being a series of human actions by citing the BS 3138 (1979) interpretation of the American Society of Mechanical Engineers (ASME) symbol for recording *Operations* as:

A definable step in a process, method or procedure. Some change usually takes place, for example, a hole is drilled, timber is cut to length, a crane is loaded.' (ibid.)

Once again, looseness of terminology and lack of a generally acceptable "building

industry" vocabulary, highlights the weakness of the CIT as failing in its objective to make good that very deficiency.

(b) The place of meaning.

The principal function of the vocabularies within this present research is to list tokens (for terms and words) that a machine will be able to match with the words used in written interpretations of maintenance tasks. This exercise lies between the first and second levels of a hierarchy at which translations can be made (summarily, syntax and semantic) and attributed to E. Niga (Nagao, 1989);

- 'The first is based upon traditional linguistics and entails the 'mapping' of the words and grammatical structures of one language...on to those of a second language...The second level is based upon communication theory, and involves the construction of sentences in the target language which have the same meaning as expressed in the source language, regardless of the degree to which the translation retains the grammatical structure of the original text. The third level is based upon sociolinguistics. That is, a sentence which is expressed within a given culture will be understood and lead to specific behaviour within that culture;..This view..is an extremely sophisticated one, in which consideration must be given to the results produced in the light of the relationship between the sentence and the cultural setting.' (p.49)

In this present research, the target and source languages were, essentially, two different sub-sets of English. Accepting Niga's definition for 'second level' translations, the question resurfaces, (again without attempting to define 'meaning') *Where is meaning expressed?* A description, once it has entered the system, re-appears in the classification, without it being *understood* by the translation process. There remains, of course, a certain amount of interaction between user and machine, none of which needs to be conducted in the target language. These might include the physical location of a document, types of repairs or processes, numbers of faults related to defined structural elements and other statistical analyses. The entire process is conducted in the same culture, therefore, establishing *meaning* is unnecessary: the computer blindly handles

symbols under programmed instruction, the user only to ensure that classing is correct. An immediate result of this conclusion, is to consider a semantically empty target language. Processing efficiency is enhanced by having terms of one size, to represent word-strings of varying length, and a suitable alphanumeric code is available in the CIT.

Reference numbers are given to each entry: J0805 [Plates], J0847 [Lapping tiles], J0851 [Single lap tiles], (Fig. 4.1 and see Appendix A) give some idea of the scope for text compression. In addition, the prefix letter signifies a table of related terms, G for 'Operations', H for 'Materials', J for 'Parts of construction works' and so on. A difficulty becomes apparent when the user is required, or wishes, to modify the translation process. For example, the vocabulary might need revising, or the veracity of a classification questioned, then a "non-word" code obscures *meaning*. Ruberg et. al. (op. cit) described the drawback, although in a different context;

'For each record, the expert included a cause (and a solution) in the description...The editor facilities for completing the problem description, were, however, inadequate and difficult to use. The alphanumeric syntax stripped contextual, spatial and ancillary information from the problem description. For building diagnosis, the data...no longer represented the actual problem at the site.....As all the fields were recoded as alphanumeric codes, it was difficult for the expert to read the descriptive evidence presented in each record.'  
(p.383)

This is more than a slight disadvantage, easily overcome by not using a coded target language, it is an indication of where the translations to be made in this present work are positioned in Niga's hierarchy. The words in the vocabulary presently under discussion, and others compiled for translating job descriptions, are intended to denote similar words, rather than 'map' them, language to language. Their meanings remain a concern only so far as to be recognisable synonyms, suitable for translated sentences defining the classifications of job descriptions. The basis for translation is now proposed.

Since the target language is a product of the source language, and the *meaning* of words and semantic relations between them is a function of mental processes entailed in reading and writing sentences, then *understanding* as a human imperative, will occur whenever appropriate words are in association.

Therefore, a string of tokens taken from a single written context and placed in conventional order, will have *meaning* when presented to a reader having the same contextual knowledge. Such an hypothesis is not predicated upon knowing directly how the meaning of words are represented in the mind. To try this would be to attempt a psychological impossibility, and further, even paradigms are unnecessary for the immediate purpose.

(c) Experienced word-definitions.

Mental models are important to psycholinguists in understanding how the mind represents *meaning* and relates perceptions of the world with language. What the subject literature, and it is considerable, largely fails to address is that constructing a model, whether it be image or proposition, relies upon introspection, as does the evaluation of the resultant theory. Pursuing this argument, Johnson-Laird (1983) criticised three psychological theories of semantics for not explaining how a mental language (*mentalese*) did relate to the world.

'This omission might be justifiable if a satisfactory account of how speakers understand the properties of intensions and the relations between them could be given without considering how they grasp extensions. In fact, this tacit assumption of the psychological autonomy of intensions turns out to be false. Unless the retrieval of referents is taken into account, it is impossible to explain the resolution of lexical ambiguities, the instantiation of words in context, the vagaries in the logical properties of spatial relations and a variety of deictic phenomena' (p. 241-2)

One result of the arguments used against claims made by Fodor (1980), Fodor, Garrett, Walker and Parkes (1980) 'that there are hardly any examples of good definition of English words' (Johnson-Laird, 1976) was to frame a general principle,

'...the closer the meaning of a word to some semantically primitive notion, the harder it will be to take its meaning to pieces and to re-express it in terms of other words. It follows, conversely that the more complex the meaning of a word, the easier it should be to define.'  
(p. 221)

A series of experiments (Johnson-Laird and Quinn, 1976) and repeated with a variety of material by others (Wykes and Johnson-Laird, 1977) (Carey, 1978) supported this prediction. Their contention that 'semantically simple words such as *move*, *have*, and *see*...have a greater diversity of meanings than semantically complex ones such as *chase*, *steal*, and *search*.' is also found in this present work. The concept is useful for identifying the verbs, and verb-phrases, likely to be used in the source, for a synonym in the target language

(d) Semantic diversity of *locating* and *clearing*.

Semantic diversity of simple words is illustrated with an analysis of two groups of verbs; those dealing with operations of *locating*, and, those with *clearing*.

*Locate* is related to *Transfer* and is displayed in the CIT as follows:

Transfer

Place = Locate = Position = Lay	(Transfer by relationship to position)
Replace (i.e. transfer to original position)	"
Displace = Dislocate	"
Slip	"
Relocate (i.e. transfer to new position)	"
Align	"
Centre	"
Orientate	"
Reorientate	"
Right = Erect	(Orientate by movement involved)
Invert	"

Reverse	(Orientate by movement involved)
Convey = Carry = Transport = Conduct = Bear	(Transfer by method)
Lift = Raise = Elevate = Hoist	(Convey by direction)
Lower = Drop	"
Impel = Propel (i.e. using force)	(Convey by method)
Blow = Blast	(Impel by agent)
Suck	"
Send = Transmit	(Transfer by material transferred)
Pour (i.e. fluids)	"

*Place* is the preferred term for *Locate* and the prefix *re* introduces meanings wider than a general sense of *again*. Thus, *Replace* and *Relocate* are given separate entries, with distinctly different meanings. Job descriptions may use *Replace* to imply *renewal* or sometimes, the substitution of one material or product for another as in:-

"Replace broken or missing slates with Eternit asbestos-free cement slates."

while *reposition* is more usually found in job descriptions than *relocate*. For example, the sentence in Fig. 3.3 reduced to: *reposition door closer*. Because the words, *position*, *place* and *locate*, can be commuted one for another, selection is arbitrary. *Lay*, the remaining synonym, does not match satisfactorily, by reason of its connotations with craft processes; *lay drains*, *lay flooring*, *lay bricks*, and so on. However, *Replace* is quite acceptable for *relay*, being polysemic concerning any action or process that means 'Remove some defective or broken product and put it back into a serviceable state.'

Then the statement:

replace = {renew, refit, refix, rebed, relay}

will act where the process is understood from what is being acted upon. Some examples are:

*replace ceiling tiles* S→V+A+N will denote;

Renew damaged ceiling tiles, or, Refix loose ceiling tiles.

As will,

*replace floor tiles* S→V+A+N for,

Rebed loose quarry tiles, or, Relay quarry tiles.

Extending this use of *replace* to other, similar cases, leads to impractical descriptions.

*hack out and reglaze window* can be reduced by word replacement and deletion to,

replace window, but not to the acceptable form of *reglaze window*.

Similarly,

*rake out joints and repoint brickwork*, S→V+ADV+N+CONJ+V+N will reduce to,

*repoint brickwork*, whereas,

*replace brickwork*, is a well-formed sentence, but a quite different task.

Transformation of sentences is another matter. *Replace panes* and *replace pointing* are only partially descriptive and need the specifics, in What? and, to Where? as in,

*replace panes* [window], and *replace pointing* [brickwork]

The noun is positioned at the end of the sentence because the prepositions *in*, *to*, are excluded from the minimum vocabulary. The adjectival use of [window] would contravene the general rule for polysemic words, (see (k), Chapter 4) preventing a grammatically acceptable, *replace window panes*, and also the slightly odd, *replace brickwork pointing*. Then, it must be asked, in this context is there any need to make a distinction between What? and Where? Using prepositions would allow,

*replace panes in window*, *replace panes to window*, *replace pointing to brickwork*,



(but not, *replace pointing in brickwork*, on the grounds of understood practice).

Then provided the identifying term is a noun, its position in the sequence will give meaning to the transformed sentence and Where? is subsumed in, What? has had work done to it. Some examples of this are,

*replace pointing* [ridge], *replace pointing* [chimney], *replace tiles* [roof], where,

S-->V+N+N

The possibility of excluding adjectives from the classification grammar is also considered later.

*Clearing* is a non-preferred term in the CIT and as shown in the next extract, is displayed in the table of relationships for 'disassembly'.

(Operations involving general concept of disassembly).

Remove = Clear	
Deprive = Starve	
Abrade = Abrase = Erode	(Remove by method)
Scrape	"
Corrode	"
Etch	"
Clean = Cleanse	(Remove by what removed)
Wash = Wet clean	"
Dry clean	"
Purify = Refine (i.e. impurities)	"
Sterilise (i.e. micro-organisms)	"
Decolour = Decolourise	"
Bleach	"
Dry	"
Dehydrate	"
Dewater	"

Classifying *Clear* as a synonym of *Remove*, does not prevent ambiguities arising from descriptions, such as in Fig. 6. where, *...clear blocked WC*, substitution makes the

phrase read, *...remove blocked WC*. Within the context of maintenance, *clear* may be used to mean *unblock* or, *clean* something specified by removing an unspecified object or substance(s). *Remove*, on the other hand, is usually (but not invariably) related to a specific article or matter. Therefore, if *clear* was denoted a synonym of *clean*, amending the table;

Clean = Cleanse = Clear (Remove by what removed)

the sentence, by substitution, would remain well-formed and read,

*...clean blocked WC.*

Similarly, this next description,

*...remove trap and clear blockage to sink in ladies toilet. Replace and test.*

by deleting conjunctions and prepositions, reduces to,

*...remove trap; clean blockage; sink; toilet. Replace; test.*

Both in this and the previous sentence, *Unblock* is the principal operation, but it is a word capable of analysis using *Remove* [an object] from [a pipe or opening] and, therefore, inadmissible. Had the description been written,

*..remove blockage from sink...*

it would only be possible to obtain,

*...clean blockage; sink...*

by assigning *Remove* as a synonym of *Clean* (Clean = Remove). Inevitably, the original sentences would then reduce to,

*...clean trap; clean blockage; sink; toilet. Replace; test.*

Relaxing the minimum vocabulary rules to admit *Unblock*, meaning *Unplug*, but only within the context of blocked pipes and openings, will overcome the plurality of; (Clean = Remove) and (Clean = Clear).

Sometimes, authors intend *Clear* to mean, 'to make tidy', rather than *Remove*, when related to a particular element or space. However, it is usual to add a qualifier, such as, *up, off, or, out*, as in this next extract,

*...clear out classroom after damage in demountable classroom and cart away in skip.* and *Clean classroom* is, most probably, the principal operation. Had the preposition been *to* and not *in*, then removing the *demountable classroom* must be equally possible. 'Cart away' is a verb-phrase meaning, 'convey to another place', *place* being the skip, not the eventual destination of its contents. The writer has, apparently, assumed that *skip* is sufficient to indicate both the extent of, and what, was conveyed. Now *Remove*, because it is expressed, must fulfil the role of denoting 'convey', and as shown here, not anything definite, but something arising from the associated activity. Leaving aside any consideration of the need to specify the conveyance, the reduced sentence is,

*...clean classroom; classroom; remove.*

The resultant sentence retains the doubt as to what has been removed, but without reference to the classroom being *demountable*, an informed reader is unlikely to infer removal of the building itself. Rephrasing the last example sentence,

*...clear classroom after damage to demountable classroom and clear away to skip.*

it then reduces to,

*...clean classroom; classroom; clean.*

showing the need for the second reference of *clear* [away] to mean *Remove* [Convey], purely on grounds of word-position. Such a differentiation is preserved in the vocabulary by assigning appropriate adverbs and prepositions to the synonyms of [Clean] and [Remove]:-

Clean = Cleanse = Clear = Clear off = Clear out = Clear down = Clear up

Remove = Clear away = Clear all

Thus, a description such as,

*Clear site*

will be translated,

*Clean site*

implying that the task includes the disposal of some undefined material, as certainly as this,

*...remove fibrous ceiling tiles. Clear all rubbish from site.*

will, by deletion and substitution, result in,

*...remove ceiling tiles. Remove rubbish; site.*

However, allowances cannot be made for every idiosyncrasy of English usage and inevitably, some descriptions for *Clean* and *Remove* will defeat any attempt to achieve the correct, or intended, translation.

(e) Other uses for *remove*.

*Replace* and *prepare* are other usages of *Remove*; it is inferred when *Replace* is the first term of a reduced sentence, and is expressed as a term signifying preparation for a prime process. Whatever words are synonymous in the latter sense, they can be comprehended when *Remove* is written as the initial activity. Two sentences from a previously cited job description, illustrate this point.

*Remove pan, replace pan connector, refit pan for test.*

*Remove cistern again, adjust and refit, test OK.*

In both cases, *Remove* signifies 'uncouple and set aside', or, 'take out for re-use' as defined by *refit*, not 'Clear away', previously considered with *Clear*. Now this duality in *Remove* is both qualified by physical distance, for example, 'take down' and 'take away', and, operational sequence. The translated sentences (by deletion and word replacement)

for removing and disposal of 'fibrous ceiling tiles' (previous page) demonstrate the manner in which word-position gives meaning, not only within a sentence, but also between sentences.

Remove as convey. Examination of the data (see Ch. 8) supports the realisation that their authors normally mean 'remove something from the site of operations', having first described the prime operation or task. Schlesinger, I M. (1995) in an exposition of the logical relations of verbs with their objects, proposed his idea of *Defining Participants* (DP) for differentiating the direct objects as '...participants inherently implicated in the event...expressed by the verb...' (p. 58). In developing this hypothesis, his illustration of 'Locative' DPs is of relevance here:

'(10)            She put the book on the table.

The lexical entry for *put* includes a DP for the thing that is moved and one for the place where it is put; it is inconceivable for putting to occur without any such "destination". The competition between these two DPs for direct object position is always decided in favour of the thing that is moved, the "destination" never being expressed on direct objects.' (p. 61)

When applied to the following description, taken *verbatim*, from an invoice for some work carried out by an electrician, 'destination' has not occurred (Note the lack of punctuation; upper case was used when typing accounts for presentation).

TO ATTEND SITE  
TEST OUT SIGNS (NOT WORKING) REMOVE  
CLEAN UP AND LEAVE TIDY.

The verb-phrase, *Test out* infers that the signs were illuminated, whilst the absence of any narrative to do with repairs after (*not working*), indicates they could not be made to work, so were removed. A search of the documents used for preparing this invoice, revealed that the signs were neon-lit, and also confirmed the final line,

*clean up and leave tidy*

as meaning, 'leave the scene of operations tidy'. Admittedly, even the electrician's

wording of his Confirmation Note (Order No. MB4629) , is not definite that 'remove' implies [convey], although being unequivocal about finding one faulty sign,

'...reversed operating u/s.sign from 2nd tran sign flickered - could not find no break in tubes or wiring disconnected flex preventing sign flickering - sign to be repaired (*sic*) or replaced Sign Stores'

A conclusion can be drawn from the final two words, given the underscoring of *Stores*, to mean, 'taking away a defective sign', but a transformation of the original wording must rely upon *Remove* (by assuming the unexpressed Defining Participant *site*) to denote 'take away signs from site', and *clean up and leave tidy* to mean 'clean site'. Comparing *Remove* in this present case, with the way it was used in the two immediately preceding examples, demonstrates the necessity for identifying its role by, (1) position in a sentence; (2) the type of succeeding terms in that sentence; (3) its relationship with any precedent sentence.

*Remove as Detach.* The appearance of *disconnect* in this extract from a description analysed previously (Fig.3.11) introduces another condition of *Remove*:

*... and disconnect electricity supply to damaged sign, take down sign and remove from site.*

Because *disconnect* is not charged with the same connotations as *unfasten*, *loosen*, and is more akin to *uncouple*, and given its specific relationship with services (electricity, gas, water) it is a vocabulary term [disconnect] = [uncouple, unclip, isolate]. However, 'unfasten', 'unfix', 'loosen', are all actions synonymous with *detaching* an object. Then, 'take down', 'take off', 'lift up' (*uplift* in Fig.3.9), 'take out', involve a degree of movement, they do not necessarily imply 'transfer of an object to another place', nor of destruction, 'doing away with' as understood in terms like, 'demolish', 'knock down', 'strip', 'smash', 'scrap'. This is yet another problem emphasising the conflict between restraints of a minimum vocabulary for translating descriptions, while retaining the

intentions expressed in the language used by their authors. The difficulty is highlighted by the final verb-phrase confirming *Remove* to mean *convey*, because 'site' is the locative Defining Participant.

A resolution is proposed for the use of *remove* as follows;

if *remove* is used by an author as the first word in a job description to denote the removal of a specific object or substance, then 'Remove' = [remove]

if *remove* is used again in the description, it's replacement is decided by context-

Services 'removing blockage' = [unblock]

All Principal divisions 'remove rubbish/waste' = [clean]

dismantle = [take down, take up, take off, unfix, loosen, unfasten]

These rules are expressed here in English for clarity, and need to be written into the computer programs in the appropriate language (Prolog, QBasic).

(f) Verb to Noun-class with the *ing* suffix.

The decision to change the verbs listed in the CIT from progressive to base form (e.g. *conveying* to *convey*), make it necessary to retain the *ing* suffix to terms that are both verb and noun. In this present analysis, *pointing* must be the noun and *point*, as a verb, will be inserted in constructions, S-->V+N; S-->V+N+N; S-->V+N+V+N. For example,

*point brickwork; point stonework chimney; remove joints point brickwork*

Johnson-Laird and Quinn's (ibid.) case for semantic diversity is well served here. The Shorter OED (1973) has over forty definitions of *point* and seven for *pointing*, compared with the verb [point] and noun [pointing].

However, not all commonly used building terms conform to *ing* = noun.

### Formless products:

#### Concrete

Verb (to [concrete]) and thus, ([concrete] [object]) should not be expected.

Adjective ([concrete] [object]) is ambiguous. and so is ([concreting] [object])

It follows that Noun [concreting] must revert to [concrete] so  $S \rightarrow V+N+N$  will accept lexical insertions such as, *replace concrete pipe; place concrete fill; repair concrete slab* and [concrete] cannot be a verb. An argument can be put forward to support the past tense *concreted* as the admitted term for an operation, and this is considered later in Chapters 6 and 7. Other formless products, [paint], [plaster] can be treated similarly, and there are other considerations: using *paintwork, plasterwork*, as nouns in some contexts; under the *ing* rule, *replace painting; painting window*, are,  $S \rightarrow V+N$  and  $S \rightarrow N+N$ ; *replace plastering*, and, *replace plaster*, become  $S \rightarrow V+N$  and  $S \rightarrow V+V$ , and finally, *paint brick wall*, is  $S \rightarrow V+V+N$ .

### Formed products of uniform size.

#### Brick.

Verb ([brick] up) Preposition inadmissible, therefore verb form must be disallowed.

Adjective ([brick] [object]) has same ambiguity as [concrete].[concreting]

Noun (A [brick]) and [bricking] will create confusion in the classification.

Similar problems arise with other formed products, for example, [tile] inserted in;

$S \rightarrow V+A+N$  gives *tile tiling roof* or *replace tiling roof* and limiting to nouns as;

$S \rightarrow V+N+N$  gives the more robust *replace roof tile* if [tile] is not a verb.



(g) The prefix *Re*.

The prefix *re* may be included in the vocabulary, albeit on a restricted basis. There are obvious advantages in having a list of terms for craft processes of the kind, [repoint] [rebed] [reflaunch] [rebench] [refit] [reglaze] [rehang] [replaster] [repaint] [rewire], all of which imply sequences of sub-operations necessary for the prime process itself. These usually begin with preparation, removing some defective material or component, cleaning or keying surfaces, mixing materials and so on, and concluded by clearing away accruing waste and rubbish. And it is these very sub-operations that prevent the general adoption of *re*, because they can also act as the prime process. A task that begins where 'preparation' had been completed as a *process*, by the CIT's definition of 'internal dynamism' (or at least, deliberate damage caused the onset) is such a case. Transforming the sentence, *removing broken glass and board up windows*, into *reglaze windows with boards*, gives a well-formed, but contradictory statement. An answer to this particular problem, is in part, provided by excluding the word *broken* from the vocabulary and then allowing the job description to determine the selection of [reglaze].

(h) Specific descriptions.

Verbs allow descriptions to be specific in their meaning by using a simple word to denote an operational sequence. *Repair door frame* is a non-specific case, in that a method or process is not described, but nevertheless, leaves no uncertainty over the fact that a door frame has been repaired. Being 'specific', such as this next fictitious description:

*Remove door, cut out and splice jamb of frame, replace door. Clean up rubbish and leave site tidy.* provokes further questions.

Does it mean the door was removed and replaced by a new one? Or was the existing

one rehung? What is different about this job compared with the previous case? Very little, if the final phrase is assumed as empty embellishment. The following description is 'specific' and also demonstrates the need to retain terms such as [Splice].

'To attend site and splice repair Infants entrance door by hinges and refix and adjust door closer.' (5671.TXT)

The hanging stile of a door has been repaired by splicing in a new length of timber, and the door closer refixed and adjusted. It is another example where the language used in the original description shapes the specificity of the task, but does not help to formulate a rule for its definition.

Returning to the sentence analysed in Figure 3.9 (Chapter 3), can excluding the verb, *polish* from the vocabulary be sustained? Appending, *grout and polish off*, to the reduced phrase (.....1.4) then gives,

*rebed quarry tiles, grout and polish off*

Here, the main task is to relay a particular ~~Type-of~~ floor tile, with a superfluous final phrase that would not be translated if *grout* and *polish* were excluded. The action of polishing is a consequence of grouting, which itself is a necessary part of of the floor tiling process. Supposing the task of grouting was invoiced separately, as,

*To attend site, grout and polish floor tiles.*

the reduction *grout floor tiles* is a satisfactory sentence; S→V+N+N (avoiding *floor* as an adjective). Matters change when polishing becomes an operation,

*sand and polish wood block floor* S→V+CONJ+V+N+N+N

Operationally, *sand* and *polish* are sequential, but can be independent tasks. The CIT classes [sanding] a synonym of [rubbing down] '(Smoothing by secondary operation)' and [polishing] the preferred term '(Smoothing by degree)'. (see Appendix A, p. VII)

This difference of emphasis in the sequential role of *sand* and *polish* is resolved by admitting both as operational terms, making [sand] the preferred term for those listed in the CIT. Further, if *sand* appears in noun form, it will be translated [sand] and not [rubbing down].

From these examinations of verb-noun-adjective words, the indications are that ambiguities cannot be overcome with general rules unless every operation and process is denoted from a very limited set of terms, such as, remove, clean, repair, demolish.

Restraints of the minimum vocabulary's principles have to be relaxed and complex words admitted, even if some are definable in terms of other words in the vocabulary. Such dispensations can be restrained by mapping operations to individual classes, such as, [weld], [solder], to SERVICES

A selection of preferred specific terms from the CIT tables, (herein Appendix A) follows:

level; sand; polish; hack; compact; align; ease; dismantle; scrape; demolish; weld; assemble; paperhang; grout; excavate; support; test; disconnect; reglaze.

Note the deletion of the *ing* suffix, discussed above (para. (f)). Synonyms are not given because they are examined again in Chapter 7, with proposals for a classification language and scheme.

### Summary

The role of verbs has been considered, particularly contradictions of the noun-verb or gerund, and rules proposed for their resolution. Further ambiguities can be expected, not only from the canons of a minimum vocabulary, but because any classification scheme will impose constraints by influencing the choice of class terms. Some of these aspects are examined in the next Chapter with a proposal for a classification grammar and scheme.

6.

A review of classification.

The previous chapter showed that the Construction Industry Thesaurus did not mutually exclude terms with the rigour essential for an automatic classification, and in order to develop the approach of limiting lexical selection, it is necessary to examine classification at a general level and evaluate some of those currently used in the building industry.

(a) Classification approaches in general

Reading the literature an Information Theory indicated several approaches for structuring a classification, the most apposite being Tomlinson (1969) who described a generalised classification, based principally upon, amongst other antecedents, a refinement by Feibleman (1954) of the Theory of Integrative Levels, which in turn, can claim origins in Auguste Comte's classification of the sciences. Briefly, the Theory postulates an arrangement of entities into a sequence of levels, whereby an entity with at least one new specific feature is produced by the organisation and incorporated into those existing on the preceding levels, whilst its function is defined in terms of higher level entities. There are twelve 'Laws of the Levels', but of greater significance are the 'Rules of Explanation':

1. The reference of any organisation must be at the lowest level which will provide sufficient explanation.
2. The reference of any organisation must be to the level which its explanation requires.
3. An organisation belongs to its highest level
4. Every organisation must be explained finally on its own level..
5. No organisation can be explained entirely in terms of a lower or higher level.

Because no classification is capable of placing all entities in a simple linear structure, it

is necessary to have a branching series of integrative levels, which as Feibleman (1954) explains,

'...that certain levels build up to two or more fields which may either subdivide further or come to a more or less abrupt end of level building.

Tomlinson (1969) cites the 'most obvious example' being where living is differentiated from non-living matter, and goes on to say,

'If entities from these two series of levels were merged into sequence, the levels could only be defined very generally, and it would be necessary to equate living cells with minerals, or a geographical feature with an organ. Clearly, the further away the levels are from a point of divergence, the further they will be removed in meaning' (p.29)

The general classification is then comprehensively described and expanded in six further papers using terms from physics, mining, geology, sculpture, and concepts within politics. Philosophically, there are defects in the scheme, which to some extent reflect the anomalies in Feibleman's Laws of the Levels and his Rules of Explanation which accompanied them. But suitably modified, it gives shape and a conceptual framework within which the building maintenance classification scheme is constructed.

The arrangement of the scheme is outlined above in Figure 6.1 Only the level sequences for Artefacts and Activities are detailed because of their particular relevance. Artefacts embrace the majority of entries and their importance merits as careful and precise an interpretation as possible. For this reason, the explanation by Tomlinson (ibid.) is quoted verbatim and follows.

'Artefacts.

Entities and substances, non-living, that do not occur naturally but are man-made or processed. A sequence of five levels has been defined. These levels show an increase of specificity of function as well as an increase of complexity and organisation

I. Raw Materials This contradicts the definition of this group as a whole in that it represents substances that do occur naturally, but there would appear to be a need for this group in that it contains materials that are the starting point for some process of manufacture by man, and the terms included at this level imply such a use by man. The terms included here may in fact be also accommodated elsewhere as physical or chemical entities. This needs further examination.

II. Worked substances: These are substances that do not occur naturally in this form, but are man-made or processed, e.g. glass, concrete, paper.

III. Components These may be natural or man-made materials that have been worked into some definite form and have a homogenous consistency, e.g. bricks, planks, concrete blocks. All these have a very wide range of applications.

IV. Assemblages: Like or unlike entities from level III assembled together in such a way that they are organised to perform a specific function, but whose definite function can only be realised in terms of higher levels, e.g. a motor car gearbox, which may be the organisation of various components, but its activity can only achieve its full purpose when the gearbox forms part of the car as a whole.

V. Finished Complex Articles These consist of a number of components or parts, and can perform their designed function in this state; the function of entities at this level generally being of a more specific nature than that of entities from lower levels.' (p.32)

The second category, 'Activities', portrayed in Figure 6.1, does not comply with the Law of the Levels. '...in general these levels have been considered as related to the...entities of which they are the activities.' (Tomlinson, 1969). Only three seem relevant. One is 'Directed Motion' because it '...includes any movement where the direction is implied and includes such terms as circulation...change of position.' (ibid). Another of these pseudo levels is 'Self Activities', 'These are the intrinsic activities of an entity which may be implicit in its definition,...' (ibid). Finally, the subdivision of 'Causation--by an agent', 'The action...on an entity by some other....' (ibid) Roberts et al.

(1970) followed a similar argument in compiling the CIT and was discussed in Chap 3 (a) (and see *Note* in Appendix A).

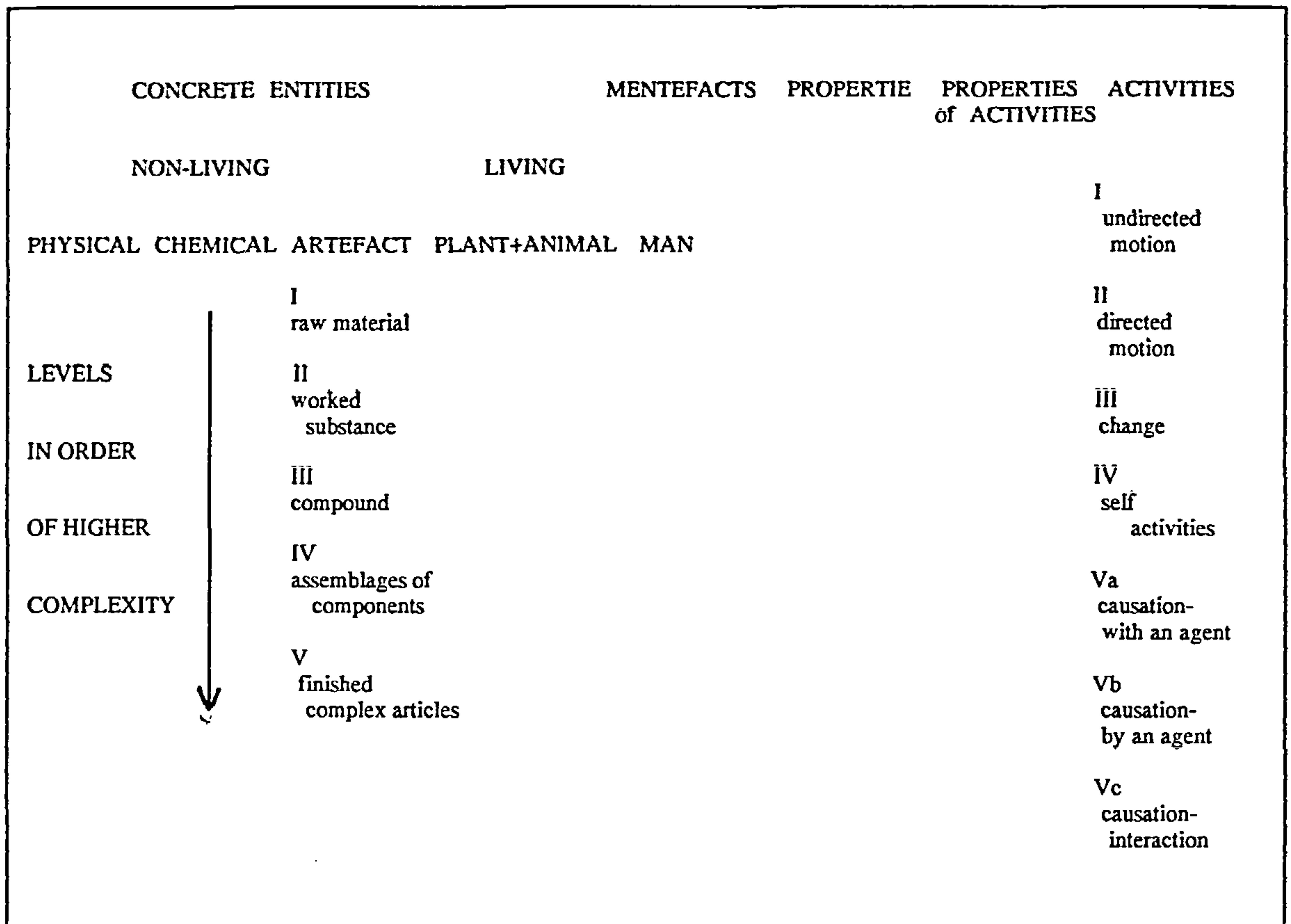


Figure 6.1 General classification scheme based upon Integrative Levels.

after Tomlinson (ibid) (Fig. II)

(b) A review of some existing classifications for construction work.

Classification and indexing systems applied to construction are, very broadly, of two kinds: pre-indexing and organisation of production information; coding and classification for pricing and cost controls, a group which includes schedules of work descriptions. Although each application has different functions, their underlying structure is a series of levels moving downwards from the complex to the simple or primitive, reversing the direction taken for the Integrative Levels just examined. The following is a review of



those influencing, in some way, the categorisation of descriptions and the objective of this present research. The first is widely known by the acronym, CI/SfB and after some 40 years is still being developed, the most recent version is 'Uniclass', to be published in 1997, (National Building Specification, 1995)

CI/SfB was developed from Samarbetskommitten for Byggnadsfrigor (SfB) a Swedish system, and the Universal Decimal Classification (UDC) by the International Council for Building Documentation (now known as CIB). The current version is administered in this country by the Royal Institute of British Architects under licence (RIBA, 1961 & 1968). Information is organised into four divisions, each represented by a distinct form of notation.

Table 0	Built environment	Symbol = numbers
Table 1	Elements	Symbol = numbers in parentheses
Table 2/3	Construction form	Symbol = letters
Table 4	Activities and requirements	Symbol = letters in parentheses

Each table is arranged as a matrix with cells denoted by their symbols. Rules of indexing are given to enable documents to be filed and later retrieved. The Built environments scheduled in Table 0, where maintenance tasks are carried out, hold no significance for job descriptions. Working in the 'cellar' (981) of 'housing with basements' (812', rather than a (981) in a 'public house, bar or tavern' (517) is only meaningful to the operative concerned. The other tables contributed to the classification presently under discussion, at least to the extent of providing class titles. This is not a criticism of CI/SfB, but an inevitable result of it's purpose for filing documents and of it's structure, which like the CIT is a faceted-classification. Vickery (1960) defined this as:

'...a schedule of standard terms to be used in the subject description of documents...the terms are first of all grouped into homogenous subject fields...Within each subject field the terms are divided into groups, known as *facets* and within each facet they may be arranged hierarchically. The facets are listed in a prescribed order, which is usually the order in which terms are to be combined to form compound subjects. By means of this combination order, relations between terms are displayed.' (p.9)

Systems based upon this concept correspond to the integrated classification already outlined. Thus, a document dealing with condensation on aluminium ceiling tiles in swimming pools would be filed according to rules (RIBA, 1968),

'...under a reference of 541 (45) Sh4 (I6) which represents the class 'swimming baths' 541, subclass 'ceiling finishes' (45), subclass 'tiles/aluminium' Sh4, subclass 'condensation' (I6)' (p.20)

But this is a special case and such a document would be defined by the dominant feature of 'condensation', rather than being perceived as a conceptual level of one problem with aluminium tiles in an individual circumstance.

Although the CI/SfB classification has a close affinity with the structure proposed by Tomlinson, (1969) it has a stronger relationship to the CIT and was used in this present research to provide class and category titles for vocabularies and a supplementary source for lexical entries.

Table 1 of CI/SfB provided class titles, for example, Primary elements; secondary elements; finishes; fixtures; (as subclasses of components) and distinctions between *services* and *installations*. Most of the facets tabled identified **Type-of** and **Part-of** terms.

Table 2/3 contributed a listing of construction forms and materials, augmenting the definitions given in the CIB Master Lists (CIB, 1972) discussed in Chapter 4 (b).

The Construction forms heading the columns of the matrix were:

cast in situ; bricks, blocks; structural units; sections, bars; tubes, pipes; wires, mesh; quilts; foils, papers (except finishing papers); foldable sheets; overlap sheets, tiles; thick coatings: rigid tiles; flexible sheets, tiles; finishing papers, fabrics; thin coatings; components; products in general.

The rows classed materials in three groupings:

Formed products;      Formless products;      Agents, chemicals.

Some terms listed under each facet of the matrix were excluded by the rules for minimum vocabularies and others contradicted definitions used earlier. One instance is the meaning of 'Component'. The CIB Master Lists regard it as any product having one, or more, expressed purpose, but the CI/SfB definition is 'Complex single purpose products ...which are manufactured off site.' (p.78) Thus, 'Specific components' are classed as 'windows, doors, trap-doors, balustrade units etc.' (p.78), while in Table 1, under 'Services' (5-), the facet 'Drainage' (52) includes, amongst several other 'Parts: inspection chambers,... pumps, gulleys,...rainwater butts...' (p.54). Further, 'Sanitary fixtures' (p.59) lists as Types, such fittings as baths, which CIB Master Lists cite as 'components'.

Table 4 (CI/SfB) has two subject or feature groups, **Activities**, and, **Requirements**. The latter was wholly inapplicable to this present research and of the former, only one row from the matrix of facets specifying 'Construction Operations' was relevant. These were:

Protecting; Clearing, preparing Transporting, lifting; Forming, cutting, shaping, fitting; Treatment, drilling, boring; Placing, laying, applying; Making good, repairing Final cleaning up. (pp.90-91)

All of these facets qualify for the pseudo level that Tomlinson (1969) designated, 'Causation --by an agent' and expanded by CI/SfB, 'Operations on site resulting in construction' and, 'Construction work, plant and operations together' (p.99). But it must be emphasised that the system, although structured on integrative levels, remains a

cataloguing syntax for storing and retrieving documents.

(c) Coding and maintenance classifications.

Classification of building maintenance is driven by the public sector's demand for more, and accurate, feedback information, to improve financial controls and exploit computer technology. Codes are based upon a rational grouping of the work, emphasising entities. Descriptions are not generated at the operational level, but by the staff managing an estate of known size and composition: an instruction might be given a classification code by referring to a data-base of property details, before the work is carried out. As well as planning and controlling repairs, it is possible to carry out statistical analyses, identifying design faults and predicting life expectancy for components and building elements. Many of the items of information generated are concerned with property management, and are therefore, not relevant to this present research.

Randall (1979) described such a system, which was introduced by the Property Services Agency (PSA) in 1973, whereby 'useful knowledge gained during the life cycle of buildings and works is collected, analysed and fed back..' (p.1). Agreement about the general dissemination of such knowledge into other domains has, according to Skinner and Kroll (1984) proved to be elusive;

'...perhaps 1000 elements will be needed for a common framework to cover various kinds of building. The situation is less clearcut with description or location. The former ought to distinguish repair from replacement, so that service life can be determined, but ability to separate accident or vandalism from wear and tear may be a matter for a particular organisation...' (p.54)

The Building Research Establishment, Princes Risborough Laboratory (PRL) commissioned the Building Research Unit, Bristol Polytechnic to carry out a research project into developing a general classification code. (Holmes, Droop and Mellor, 1985),

(Holmes, 1987) A standard format was being established, yet the authors noted that many different coding systems had been devised, representing,

'...the viewpoint of the individual that commissioned it and the compromises that he, or she, is prepared to make.' (p.2)

Five years later, describing a computerised maintenance management system in the housing department of one London Borough, Saville (1990) produced supporting evidence of this idiosyncrasy. With a building element code of forty items, and a schedule of rates compiled by the Authority's staff, the scheme was more detailed than the general classification for PRL, yet provided very similar information.

Two case studies, one based on 370 houses and the other covering 50 primary and 10 secondary schools (Holmes and Mellor, 1985) demonstrated the flexibility and strength of the Bristol code, and that it was capable of providing fine detail, with potential for expansion, should it be required.

Their coding system is made up of six separate parts, building a total of fifteen digits;

an heirarchical code (6 digits) defining the element, secondary element or component requiring maintenance;

a process code (1 digit) representing the operation;

a descriptive code (2 digits) for detailing materials and/or location of selected elements and components (doors, windows, fences, gates);

a block code (2 digits) to identify individual buildings, or their parts, in any establishment;

a location (house) code (2 digits) '...probably limited to the initial repair instruction.' (Holmes, Droop and Mellor, 1985 p.4);

a reason code (2 digits) for identifying the type of maintenance (planned, unplanned) and, where possible, the causative agent, a difficulty recognised by Skinner and Kroll, (1984).

The first two parts make up a job code and are the only ones of immediate interest.

This elemental code starts with nine primary groups:

1. External painting.
2. Internal painting.
3. Structure.
- 4 Structural fixings and internal finishes.
5. Plumbing (excluding Heating).
6. Heating and other services.
7. External site works.
8. Ancillary services.
9. Other buildings etc.

These are further divided and sub-divided into a hierarchy with five levels (in addition to the primary group level shown above) allowing a significant range of elements to be included.

One example, taken from 'Plumbing (excluding Heating)' demonstrates how a code is developed and the minutia it can contain. Commencing with the first column (level One of the Elemental code) in Figure 6.2, '5' denotes 'Plumbing (excl. heating)' then moving to the right, further digits are added until the lowest, and most particular level gives a code for the smallest, practical part. There are four ball valves in the fifth level tabled against '52 Water storage' (5 + 2) and the chronology of code building for

- 5202 = Unspec. storage + Fittings + 1 = 52021 Ball valve
- 5212 = Cold water tank + Fittings + 1 = 52121 Ball valve
- 5232 = Combined cylinder + Fittings + 1 = 52321 Ball valve
- 5242 = W C cistern + Fittings + 1 = 52421 Ball valve

then by adding a further '1' to the end of each of these chains, gives the six element codes for 'washer'. It should be noted that one effect of sequencing digits, in this

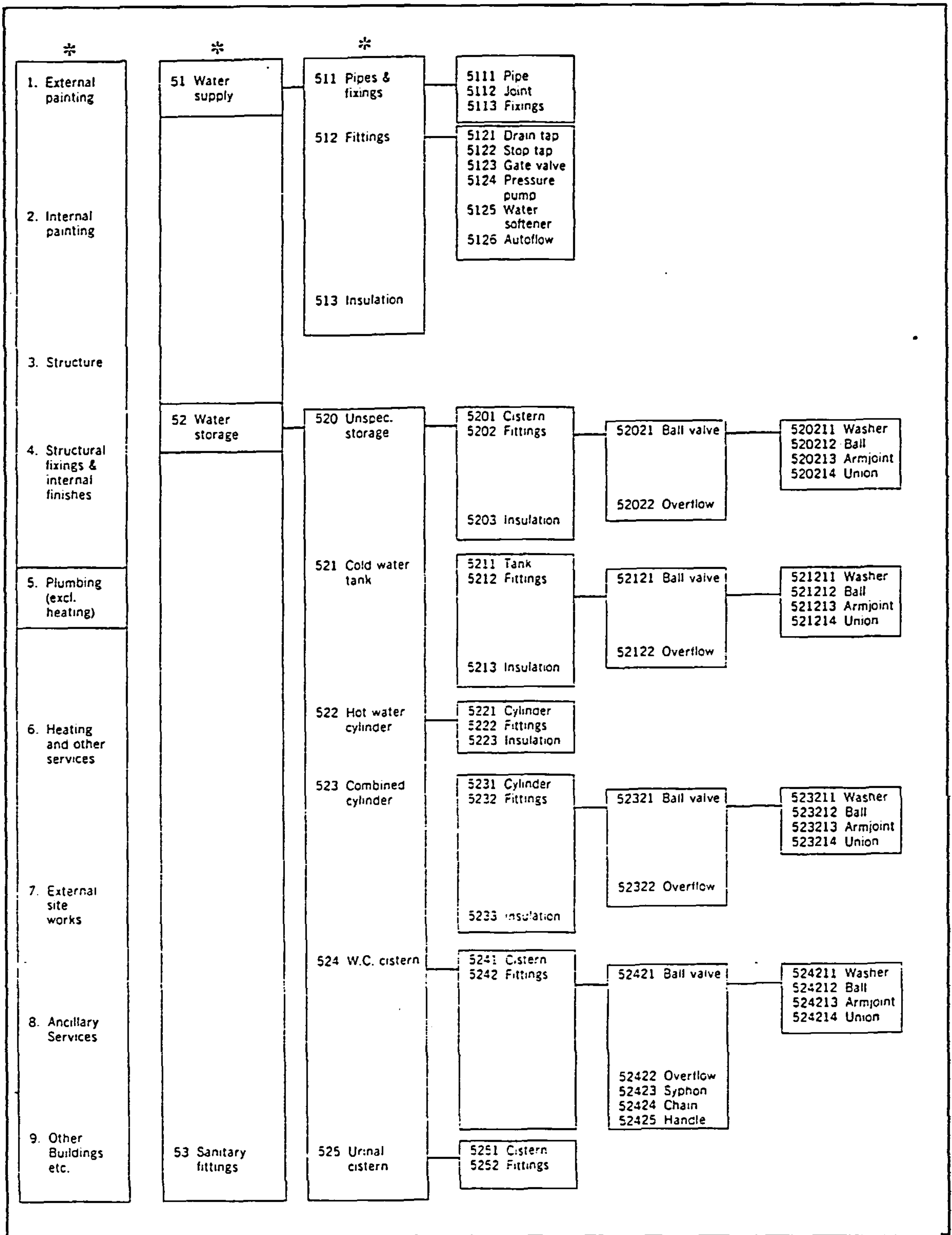


Figure 6.2 Elemental Coding (after Holmes, Droop and Mellor, 1985).

instance, is to give four different codes for an item likely to be identical. The authors counter, saying that not every organisation would need to use this level of the elemental code, and

'...would normally prefer to code 'repair to ball valve washers of wc cistern' (524211) more generally as repair to 'wc cistern fittings' (5242), or even repairs to 'wc cistern' (524). That is, they would use only the suggested truncation (asterisked in the Appendix)' (p.9)

The job code is completed with the process code, of which there are nine categories, and enable an operative to be instructed to carry out the work:

1. General repair or unspecified job
2. Renew, replace
3. Install, fit (new i.e. for the first time)
4. Remove, demolish
5. Refix
6. Ease and adjust
7. Clear out, clean
8. Test, check
9. Supply only

Comparing what are effectively eight processes, with the number of operations (verbs) judged necessary in Chapter 5 (h), illustrates the reductions in vocabulary made possible by being in the position of initiating work. The certainty of knowing what kind of cistern was fitted in a particular house; a confidence resulting from access to estate records, condition surveys and the like, should not need a code that can pinpoint the location of a washer, yet is incapable of distinguishing whether that item has some significant feature. Nevertheless, an heirarchical approach, over a limited number of levels, was found to have advantages in this present research.



(d) Schedules of rates.

Schedules for pricing and cost controls, are to a large extent, aligned with construction process and the systems used for it's classification. This is a continually developing area with many influences upon its scope and direction. The classification widely adopted in this country is known by the acronym, CAWS (Common Arrangement of Work Sections for Building Works) and produced by the Co-ordinating Committee for Project Information (CCPI, 1987). Ordering generally follows the sequence of construction and has twenty-four Group headings at the highest level, using a notation of upper case letters:

A Preliminaries and general conditions

B Complete buildings

C Demolition/alteration/renovation

D Groundwork

E In situ concrete/large precast concrete

F Masonry

...ending with; Z Building fabric reference specification

There are two sub-levels, normally symbolised by ordered sets of digits.

The concept of the CAWS is a model of trade, specialist and subcontracting processes, defined in appropriate clumps of technical information, established by research and consensus to represent the technical information necessary for building design and construction. Even as a framework for studying and documenting best practice in construction information technology, there is no relationship with CI/SfB. From correspondence with NBS Services, currently responsible for developing this last

classification, (Cann, 1997) the two systems will continue to be separate entities 'since it is expected that people will tend to use different Uniclass tables for different purposes...'. This is not surprising, because the CAWS is neither a catalogue nor an information retrieval system. It does not function on integrative levels, but is a matrix of integrated processes, perceived as sets of knowledge needed to specify, draw and price work 'packages' not necessarily any particular structural element.

Thus, CAWS in satisfying its declared aims for both new building projects and alterations (per Group C) would seem to bear little affinity with *ad hoc* building repairs. Nevertheless, it was adopted by the National Schedule of Rates (NSR Management, 1995) following, with some compromises in Levels 2 and 3, the Standard Method of Measurement 7th edition, (SMM7) (RICS/BEC, 1988) in abandoning the traditional 'order of trades that had been first incorporated into rules for measuring building work in 1922 (SMM7, 1988, p.10).

The National Schedule of Rates was devised jointly between the Society of Chief Quantity Surveyors in Local Government and the National Federation of Building Trades Employers (now Building Employers Confederation) in 1982. Originally, it was confined to local authority housing maintenance, but the final publication included descriptions and rates for minor works and house improvements. Some critics thought this addition would weaken its acceptance. As Elliott (1982) pointed out in an examination of schedules and price books;

'...it is difficult to achieve a balance between comprehensive coverage that may result in an unwieldy schedule, and insufficient detail that may result in expensive operation due to 'rogue' items.' (p. 3).

Currently, the National Schedule is designed to be used in measured term contracts by public and private building owners, and is maintained by NSR Management with quarterly pricing updates. In spite of its antecedents and connections to the CAWS the

Schedule, is a listing of standard job descriptions and prices.

Summary.

The classifications and indexes reviewed here, achieve their intended purposes for cataloguing technical literature; maintenance coding systems, together with schedules of rates, meet the requirements of repair management in <sup>K</sup>profitable, controlled situations.

However, success in their respective fields make them unsuitable for automatically classing naturally occurring data, and a more appropriate building maintenance classification is developed in the next Chapter.

7.

A general classification structure for  
building maintenance identified and outlined.

The corollary to automatically classifying descriptions is to have a general classification, and of those so far outlined, none would satisfy the constraints of a minimum vocabulary or allow the freedom to accept descriptions occurring naturally from unplanned sources.

(a) Rationale for a new classification.

Willmott Dixon Maintenance Ltd., the collaborating organisation, is organised to respond to demands from their clients, but cannot forecast the range and content of instructions received (discussed further in Chapter 8). What follows is based upon the Theory of Integrative Levels, and in its train the elemental approach suggested by Holmes, Droop and Mellor (op. cit.). CI/SfB and CIB definitions are adopted wherever appropriate. Instead of numeric coding, class headings and names of specific subjects are used, which, as Sharp (1965) reported, was rare even at that time. But it would be perverse to reject an indexing method on the basis of it being old-fashioned, particularly when the string handling capabilities of present-day computer languages offer so much flexibility. To press this point, the alphanumerics identifying CIT terms, considered in Chapter 5 (b) for a semantically empty language, permit fixed field widths, vital in overcoming limitations on word-processing with Hollerith punched card input and programs written in Fortran IV (Hague,1977); employed in this present proposal, they would be an encumbrance.

Regardless of how generalized the structure of a coding system may be, it will be necessary to make additions and changes. Numeric indexes leave gaps in the number sequences in anticipation of expansion, but these either do not match the positions where

infilling is needed, or, eventually, all the numbers are used. The CIT has a potential capacity for 9999 entries in each of 10 tables denoted by a letter; the 'Bristol' general coding system Holmes, Droop and Mellor (1985) are hopeful that users will find the commonly mentioned components and:

'...any unlisted components can easily be accommodated, with their position in the hierarchy in the predetermined and all but a few categories having ample space for additions (a maximum of 9 elements in each box).' (p. 4)

Names for class headings avoid these problems. The classification always matches the entries; if a class is empty, that part of the schedule is not displayed and new categories can be added without having to revise any existing table or listing.

(b) Arrangement of the proposed classification.

There are three principal divisions

**Building Zones** corresponding to the structural elements of a building: Structures, Floors, Roofs.

**Services** including power, disposal and transport (lifts),.

**External Areas** includes spaces adjoining buildings, designated spaces unconnected with any building; operations understood to be 'External Works',

each organised into level sequences decreasing downwards according to the complexity of their contents. That is, reversing the direction taken by Tomlinson (op. cit), while retaining her interpretations (see Chapter 6 and Fig. 6.1). Level 'V. Finished Complex Articles:' has no equivalent, the closest analogy would be Place, one of the CIT facets excluded under minimum vocabulary rules. (Chapter 4 (a)).

The scheme extends for two level sequences; Level 1 corresponding directly with 'IV. Assemblages:', and, Level 2 equating with 'III. Components:'. Further sequences are not needed because 'II. Worked Substances:', and, 'I. Raw Materials:', together with sub-components (Types-of and Parts-of terms) merge to produce Level 2. They also explain

its organisation, conforming with the fourth Rule of Explanation (Fiebleman, 1954) (See Chapter 6, para. (a)).

Both levels branch into classes, accompanied by membership definitions and brief explanations. An extract from the principal division **SERVICES** is shown in Fig. 7.1. Relationships between categories, levels, classes and branches are next outlined, and the complete framework is included in Appendix C.

<b><u>SERVICES.</u></b>		
<b>DRAINAGE</b>		
<b><u>Level 1</u></b>	<b>Above ground</b>	<b>Below ground</b>
<b>Explanation:</b> Systems by the material carried and method. Chambers for access, collection, treatment.		
<b><u>Level 2</u></b>	<b>Components</b>	
	<b>Parts-of further classed by Types-of material (ceramic, plastic, cast iron)</b>	
<b>Explanation.</b> Parts; normally, formed products as purchased; fittings to smallest sub-component; formless products (trench fills, surrounds and support).		

Figure 7.1

Extract from proposed classification

### (i) Building Zones

The term 'Zone' was chosen to symbolize loci of operations, rather than expecting authors to associate their descriptions with the concept of primary elements performing one, or more, functional requirements; strength and stability, sound insulation, fire resistance, and the like. (Building Research Station, 1959). Although the zones, **Structure**, **Floors** and **Roofs**, have intrinsic class memberships, they are extrinsically different categories, so that work descriptions can be classed with a reasonable expectation of certainty. Roof boarding and floor boarding involve identical, or similar materials and

repair processes, but are members of unmistakably different groups; a chimney breast is part of a structural support system while a chimney stack is not, and can therefore be classed under **Roofs**.

At Level 1, zones branch into classes denoting primary structural support and division of space; excepting **Roofs** that are defined 'types by slope'. Since classing an entity turns on how it has been described, an inference might have to be drawn if a specific word has not been used. For example, *loft, ridge* = pitched (roof), and, the sentence *sweep leaves from roof* = flat (roof). Conversely, it is unnecessary to seek inferences of load-bearing functions for partitions and walls. In addition to providing class headings, the primary elements accommodate general operations; *Cleaning floor; Inspecting roof;* are two fictitious examples.

The classes designated at Level 2 are groupings of objects existing as integral **Parts-of** a primary element, or requiring the 'support' of a structure. Summarising, the class for 'Sub-elements' and 'Supported' are replaced in **ROOFS** by, respectively, 'Structure' and 'Weatherproofing' to conform with practical understanding of their roles. In **Floors**, 'Stairs' are considered a distinct class of component rather than being subordinated to 'Structure', which would cause anomalies between the structure of a staircase and sub-elements, components and parts (where they can be separated and become purchasable items)

## (ii) Services

The organisation of levels follows the same pattern as for **Building Zones**, although generally a simpler arrangement, with fewer classes in Level 2 sequences. However, tension exists in the division between **Space heating** and other categories, particularly



**Water supply and Gas.** Confusion in classing became apparent when, during manual trials, no heading could be found for solid fuel appliances, particularly convection heaters that are fitted and serviced by trades other than plumbers. Even with the arrangement proposed, there would be difficulty if an Aga cooker was the subject of a job description.

(iii) External spaces

The compilation of this division was prompted by a small number of tasks found when transcribing job descriptions for the data-base (see Chapter 8) as well as completeness of the proposed classification. Again, difficulty arose in making unequivocal separations between some classes. For example, having a category **Dividers** should include a class 'Hedges', but these are more firmly associated with **Landscape**, semantically and by common use. In the event of the term [hedge] needing to be classed, *hedge* was included in the lexicon as a group in the class 'Cultivation'.

(iv) Scope.

The range of classes is limited in scope by restricting the duration of tasks to approximately two days. Extended operations of the type regarded as renovation and refurbishment will be precluded by this simple limitation, but the classification has potential for refinement if time is brought in as a parameter to group jobs within classes according to their durations. Several techniques could be applied, comparative and category estimating being the most promising (Hague, 1977)

Extending the categories will be a consequence of putting the classification into practice. Of the three principal divisions, **Services** is one capable of considerable modification,

simply by inserting the CI/SfB tables for 'Services', 'Installations' and 'Fixtures'. Several additional classes, if not level sequences might be be justifiable..

(c) Classification grammar.

Classing is carried out in stages. When a job description is taken from the data-base, it is scanned for a range of key terms, and allocated to one of the principal division, or it is rejected. Next, the accepted job is identified with an appropriate class and the classification grammar rules applied in a dual role. First, a fresh heading is written, if required. Second, the job description is transformed, or reduced to verify whether it is:

a match with previously classed entries,

an analogous task and therefore, another member of that class or category,

the initial entry under a newly written heading,

a failed classing (*a priori* True or False).

This process could, superficially, be regarded as meeting the definition for the 'best' rule, stated by James (1985);

'...the overwhelming majority of applications make use of the "obvious" criterion of minimum error...the classification rule we seek is optimum in the sense that minimises the "total error of classification" of TEC.' (p.7)

In addition, the pertinancy factor proposed by Sharp (op. cit.) might also be satisfied to unity, it being the ratio between the number of relevant documents retrieved and the total retrieved. If group headings are generated automatically, from the entities being classified, and then the entities themselves are placed in those class-groups, can the TEC be other than 100%?

One answer is a qualified 'No'. Given unambiguos sentences, such as, '*glazing window*'; '*re-wiring power point*'; '*meeting Mr Jones*'; it would be surprising if the first two were

not entered, respectively, under, 'Building Zone: Structure; 'Services: Electric'; and the third was accepted. Nevertheless, classing would also, in logic, be successful if all three were rejected because headings had not been established. In practice, sentences are seldom so clear cut, and no computer program, written in a declarative language, will run in quite the way it is expected ('intended' is an attribute of procedural programs). Only manually classifying the same descriptions will allow an arithmetical comparison. But a different result will raise further questions. Are the terms in the data-base replicas of those in the job descriptions? Has the classifier corrected misspelt words?'. In short, was the test fair? Satisfying all of these doubts will neither confirm, nor deny, scientifically, the proposed classification scheme's effectiveness. That can only be measured empirically and the result positioned subjectively on a qualitative scale, which begs the question.

James (1985) applied his test to mathematical classing, whereas entries to the scheme under discussion are self-evident, which Richmond (1965) suggested, might not exist 'to serve as a basis for axioms.' (p.36). The issue is analogous with Russell's (op cit) examination of qualities. After explaining three different, physical, ways of defining the colour 'red', he pointed to their shared artificial precision and argued,

'The word "red", like the word "bald", is one which has a meaning that is vague at the edges. Most people would admit that, if a man is not bald, the loss of one hair will not make him so; it follows by mathematical induction that the loss of all his hairs will not make him so, which is absurd. Similarly, if a shade of colour is red, a very tiny change will not make it cease to be red, from which it follows that all shades of colour are red.' (p.276)

Therefore, classification errors in this present research can be assessed in two stages. Initially, to see if irrelevant jobs had been rejected. Secondly, by taking samples of classed job descriptions and then comparing their entries with a manual classing.

(d) Classification grammar compared with SMM7.

The proposed grammar is, syntactically, an ordered word-class sequence, and in that respect, is comparable with a style of building work description rooted in developments of computer-aided systems for preparing bills of quantities; a praxis officially recognised when the Joint Committee for SMM7 decided to translate the rules of measurement from traditional prose into classification tables. (RICS/BEC, 1988) Having abandoned trades order for the CAWS, It was careful to point out;

'This change makes the use of the rules a... more systematic task than the interpretation of a prose version and readily lends itself to the use of standard phraseology and computerisation. The change however does not inhibit the use of traditional prose in the writing of bills of quantities if so desired.' (p. 9)

Measurement of building works is the process of quantifying and describing the constituent parts of some planned construction, in order to establish a contractual price, and therefore, converse to the object descriptions in this present research. The distinctions marking points of dissimilarity in the 'sentences' generated under SMM7

Rules being:

written by trained personnel prior to carrying out the construction

in principle, more concerned with defining what is to be built, rather than any operational method; details of workmanship, materials and finishes may be given separately, generally by referring to British Standards and Codes of Practice;

structured so that features likely to significantly affect cost is included, either within the Measurement Rules, or because the measurer believes it would.

Measured descriptions follow Classification Tables, supplemented by 'Measurement Rules' 'Definition Rules' 'Coverage Rules', and, 'Supplementary Information' and an extract from Work Chapter F Masonry is shown in Fig. 7.2. Reference coding starts with the Work Chapter letter, followed by categories denoted with double digits and headings for construction work, when covered by the same Rules. The groupings accord with bricklaying and masonry trade work: F10 Brick/Block walling and F11 Glass block

**F10 Brick/Block walling**  
**F11 Glass block walling**

INFORMATION PROVIDED	MEASUREMENT RULES	DEFINITION RULES	COVERAGE RULES	SUPPLEMENTARY INFORMATION
<p>P1 The following information is shown either on location drawings under A Preliminaries/General conditions or on further drawings which accompany the bills of quantities:</p> <p>(a) Plans of each floor level and principal sections showing the position of and the materials used in the walls</p> <p>(b) External elevations showing the materials used</p>	<p>M1 Brickwork and blockwork unless otherwise stated are measured on the centre line of the material</p> <p>M2 No deductions are made for the following:</p> <p>(a) voids <math>\leq 0.10 \text{ m}^2</math></p> <p>(b) flues, lined flues and flue blocks where voids and work displaced are together <math>\leq 0.25 \text{ m}^2</math></p> <p>M3 Deductions for string courses, lintels, sills, plates and the like are measured as regards height to the extent only of full brick or block courses displaced and as regards depth to the extent only of full half brick beds displaced</p> <p>M4 Curved work is so described with the radii stated</p>	<p>D1 Thickness stated is nominal thickness unless defined otherwise below</p> <p>D2 Facework is any work in bricks or blocks finished fair</p> <p>D3 Work is deemed vertical unless otherwise described</p> <p>D4 Walls include skins of hollow walls</p>	<p>C1 Brickwork and blockwork are deemed to include:</p> <p>(a) extra materials for curved work</p> <p>(b) all rough and fair cutting</p> <p>(c) forming rough and fair grooves, throats, mortices, chases, rebates and holes, stops and mitres</p> <p>(d) taking out joints to form a key</p> <p>(e) labours in eaves filling</p> <p>(f) labours in returns, ends and angles</p> <p>(g) centering</p>	<p>S1 Kind, quality and size of bricks or blocks</p> <p>S2 Type of bond</p> <p>S3 Composition and mix of mortar</p> <p>S4 Type of pointing</p> <p>S5 Method of cutting where not at the discretion of the Contractor</p>
<p><b>CLASSIFICATION TABLE</b></p>	<p>1 Walls</p> <p>2 Isolated piers</p> <p>3 Isolated casings</p> <p>4 Chimney stacks</p>	<p>1 Thickness stated</p> <p>2 Facework one side, thickness stated</p> <p>3 Facework both sides, thickness stated</p>	<p>1 Vertical</p> <p>2 Battering</p> <p>3 Tapering, one side</p> <p>4 Tapering, both sides</p>	<p>1 Building against other work</p> <p>2 Bonding to other work</p> <p>3 Used as formwork, details of temporary strutting stated</p> <p>4 Building overhead</p>
	<p>M5 Building against other work and bonding to other work is measured where the other work is existing or consists of a differing material</p>	<p>D5 Battering walls are sloping walls with parallel sides</p> <p>D6 Tapering walls are walls of diminishing thickness</p> <p>D7 Thickness stated for tapering walls is mean thickness</p> <p>D8 Isolated piers are isolated walls whose length on plan is 5 four times their thickness, except where caused by openings</p>	<p>C2 Brickwork and blockwork bonded to another material is deemed to include extra material for bonding</p>	

Figure 7.2 Extract from SMM7- F MASONRY (Acknowledgements to BEC/RICS Joint Committee)

walling; F20 Natural stone rubble walling, F21 Natural stone ashlar walling and, F22 Cast stone walling/dressing: F30 Accessories/Sundry items for brick/block/stone walling. Classification Table references are cited, when required, in order (excluding the column for units). For example, 'Walls: building overhand' = F10.1.0.0.4; 'Isolated chimney shafts and the like (Nr): building from outside scaffolding' = F10.7.0.0.1. 'Type of pointing' (SMM7, F10, S4) includes a variety of finishes but they are executed 'As the work proceeds' or 'As a separate operation'.

Whether the measured descriptions are written in prose style, or strictly according to Classification Tables, verb-phrases have a lesser role to play than noun-phrases, a feature demonstrated with two measured items of brickwork, using the SMM7 extract (Fig.7.2).

Walls: one brick thick (215): common bricks (BS3921) cement mortar (1:4) (F10.1.1.1.0 S1, S3)

Walls one brick thick built in common bricks (BS 3921) laid in cement mortar (1:4)

Walls: facework one side, half brick thick (103): building against other work: Radcliffe facings PC £386.00 per 1000 delivered: stretcher bond: gauged mortar (1:1:6): pointing weather struck joint as work proceeds (F10.1.2.1.1 S1, S2, S3, S4)

Walls half brick thick one side built against existing brickwork, in Radcliffe facings PC £386.00 per 1000 delivered, laid stretcher bond in gauged mortar (1:1:6), fair faced other side and pointed with a weather struck joint as the work proceeds.

There is little to choose between the styles, either in number of words or verb-phrases. In fact, presenting these descriptions as a bill of quantities, under headings, erases most of the differences:

## Brick/Block walling

### Common brickwork (BS 3921); cement mortar (1:3)

#### Walls

One brick thick (215)

(Prose style exactly the same)

### Radcliffe facings PC £386.00 per 1000; gauged mortar (1:1:6)

#### Walls

Half brick thick (103): facework one side: stretcher bond: building against other work: pointing weather struck joint as work proceeds

Half brick thick one side built against existing brickwork, laid to stretcher bond and fair faced other side and pointed with a weather struck joint as the work proceeds.

#### (e) Syntax of 'Heading stage' grammar.

The classification grammar for the class and category headings, is required to reproduce the bill of quantities format, using headings from words matched with, or inferred from, the presented descriptions. Since objects form the classification, verbs are only necessary for classing and denoting actions, processes and operations carried out upon them. Effectively, this bisection leaves a grammar with sentences comprising nouns and adjectives, two prepositions, (Above, Below), and the conjunction 'and':

$S \rightarrow N$ ;  $S \rightarrow A+N$ ;  $S \rightarrow N+CONJ+N$ ;  $S \rightarrow PREP+N$ .

Now bearing in mind that words in the descriptions are being matched, at the level posited by E. Niga in his heirarchy of translation (Nagao, 1989) and discussed in Chapter 5 (a), then it is logical to dispense with adjectives and define terms as belonging to a word-class of nouns, but separated only for the computing programs into Type-of and Part-of classes.

The lexicon of headings for the proposed scheme in its present form, is very brief, and

the number of sentences the grammar can produce from it, is finite. Adding fresh terms might introduce new classes and categories, but lexical terms will be inserted in the same sequential syntax. Therefore, any such 'headings' grammar with a lexicon based upon the principles of a minimum vocabulary, will be finite and not a 'natural' language. (Any amendment will certainly compel significant changes to the supporting computer programs). Emphasising this argument, and to illustrate some syntactical features the complete lexicon is listed below.

## Building zones

### STRUCTURES

Walls, Partitions, Foundations, Sub-elements, Supported, Cladding, Finish and Treatment, Components, Window, Door

### FLOORS

Floor, Slab, Ceiling, Stairs, Staircase, Balustrade, 'Stair well', Access

### ROOFS

Pitched, Flat, Support, 'Chimney stack', Weatherproofings, Light, Edge, 'Rainwater disposal'

## Services

### DRAINAGE

Above, below, ground

### WATER SUPPLY

Service, Distribution, Pipework, Storage, Sanitary, Heating

### ELECTRICITY

Installations, Wiring, Control, Outlet

GAS (No additional terms)

### SPACE



Producer, Distribution, Delivery

## External Spaces

### 'HARD SURFACES'

Trafficked (= adjective), Surfacing, Pedestrian, Block/slab, Compacted, Applied, Fixture, Feature

### LANDSCAPE

Recreation, Cultivation, Natural, Shaped, Bulk, Surface, Single Group, Fence, Information.

Upper case and bold fonts are used for clarity and are not part of the grammar or computer programs. Syntax takes a hierarchical or precedent ordering, where the headings (sentences  $S \rightarrow$  as shown) are subordinated to the Level above, but if used for the first time, will generate the superordinated headings. For example, if 'Wiring' is matched from a job description and there are no other jobs in that category, Level 2 Installation, is placed first, preceded if necessary, by ELECTRICITY, and finally, the Principal division, SERVICES. In this way, the classification is protected from repetitive, and redundant, headings. Plurals take precedence in the hierarchies, as Components to Window, Door, and, Light, Edge, Rainwater disposal. Similarly there is STRUCTURES and Structure at Level 2 under ROOFS.

The lexicon of the source language necessary for matching and making logical inferences is not shown here, but an example can be seen in the program listings in Appendix E

#### (f) Verb-class patterns.

The classification grammar that will transform, or reduce sentences, has already been examined and the principal structural characteristics anticipated. Verbs make up its most significant feature. Authors of job descriptions, in effect, write;

'I went to [location] and did this [action] to this [object] then another [action] as a result of the first [action] and finally this [action]'

(and in this present author's experience, this habit is not confined to Willmott Dixon Maintenance Ltd.).

'To attend site' is the prefix to every invoice and stamps 'The work carried out at this place was...' onto the description. Unfortunately, the word *site* is also used to indicate that an [action] occurred at an undefined location. For example, *clear site, remove from site* carrying the risk of confusion with an area in the Principal Division EXTERNAL SPACES. A resolution for the shifts in meaning of *clear* and *remove* was put forward in in Chapter 5 (a), nevertheless, it wil not be possible to class this sentence:

*To attend site and clear away rubbish.*

Deleting 'To attend site' and 'away', leaves 'clear', 'rubbish' and applying the rules for selection the transformed sentence is [clean] [ ? ]. *Rubbish* is not a material and has no place in the classification, there is no location for the remaining term, or even the complete sentence *clean rubbish*. Trivial though it may seem, a weakness is exposed that is a far from trivial matter to solve as a general principle

Ambiguities such as these abound and in previous discussion, the approach of strictly limiting the choice of words has showed the feasibility of overcoming many of these problems, but not as extending to every example of idiosyncratic English usage.

From the previous analyses, several recurrent sentence patterns are evident:

S→V+N	Process + object
S→V+A+N	Process + Type-of + Object
S→V+V+N	Process + Process + Object

Examples of the last pattern are the insertions, 'Remove rehang door', Remove rehang door', Remove ease window'.

Restating the resolution for *remove* stated in Chapter 5:

If *remove* is the first word in a description then, 'remove' = [remove].

If *remove* is used again, then 'remove rubbish/waste' = [clean].

underlines the argument for including 'rubbish/waste' in the classification. This would solve this specific case:

where *S* = *take down wall and remove rubble* and is reduced to:

$S \rightarrow V+N+V+N$        $S = [\text{dismantle}] + [\text{wall}] + [\text{clean}] + [\text{material/away}]$

Under the same resolution, 'removing blockage' = [unblock] must apply to any pipe or similar hollow section, not only those classed in SERVICES. This will become apparent in the following paragraph.

(g) Vocabulary for the classification grammar.

Some terms in the vocabulary have meanings that are influenced by the context of the sentences where they appear, and for this reason, the explanation is made in the context of the class ROOFS in the Principal Division, BUILDING ZONES.

Level 1

The major processing function at this level is to infer, from the job description, whether the roof is Pitched or Flat, and is carried out by the heading grammar. Entries classed here will be those not associated with a named entity. Listed here are the verbs for every category under ROOFS.

Action processes;

[investigate] = inspect, check, examine, examin, ecsamin, survey

[clean] = sweep, 'wash down',

[replace] = refix, renew, clip, nail

[treat] = spray, apply, inject, 'brush on'

[unblock] = remove (blockage), unseal, unstop

[dismantle] = 'take down', demolish, lower

## Level 2

### Structure

#### Support (Parts-of)

[timbers] = rafter, collar, joist, joyce, binder,tie, hip, valley, flank, beam, plate, trimmer, truss, purlin

#### Chimney stack (Parts-of)

[brickwork] = oversail, stack, course

[flaunching]

[pot] = cowl, liner

[flue] = parging, lining, 'smoke pipe'

#### Type-of material.

[brick] = facing, engineering, blue, staffordshire, red, accrington, stock, fletton

[stone] = ashlar, ashler, 'art stone', 'cast-stone',

[concrete] = 'in situ', rc, RC, reinf, reinforced, renforsed

[material/waste] = rubbish, rubble, leaves, leafs, old

### Weatherproofings

#### Type-of material

[tile] = interlocking, marley, rosemary, redland, roman, pan, pantile

[slate]

[flexible] = felt, mineral, sarking, polythene, underfelt, slaters

[asphalt] = asphalte, ashfelt, asphelt

[bitumen] = pitch, 'hot pitch', 'cold-tar', black

[sealant] = mastic, waterproofer

[battens]

## Part-of

[sheeting]

[flashing]

[valley]

[soaker]

[ridge] = hogback, mono, apex

[hip] = bonnet

[boarding] = plywood, chipboard, 'T & G', chip, 'felt faced'

## Components

Light (Note: defined by words in description, such as 'dormer', 'lantern' = Heading).

## Part-of

[sash] = 'glazing bar', stile, style, 'bottom rail', 'bottom rail', 'top rail'

[cill]

[frame] = jamb, head, transome, transom, mullion, stop, rebate

[hinge]

[fastening] = stay, pin

## Type-of

[glazing] = glass, putty, pane, wired, cast, GWC

## Edge

## Part-of

[fascia] = fascia, faicher, facier

[soffit] = soffite, undereaves

[eaves] = eve, eves

[verge] = barge

[abutment] = skirting, apron, stepped

## Rainwater disposal

### Parts-of

[gutter] = ogee, OG, 'half-round'

[fitting] = 'union bracket', angle, 'stop end', 'running outlet', outlet, spout, 'stop end',

[clip]

[head] = hopper, chute

[pipe] = 'down-pipe', drop, 'swan'neck', offset

[seal] = gasket, neoprene

### Type-of material

[cast-iron]

[plastic] = PVC, UPVC, marley, hepworth

[galvanised]

[aluminium]

[asbestos] = cement

[wood] = softwood, deal, hardwood, mahogany

This listing is not exhaustive, but it includes all the misspelt words noted during the transcription of invoices to the data base. The complete vocabulary would have some repetition of terms, (shown in square brackets) and this is a consequence of using terms in a context.

## Summary

This Chapter has proposed and developed the basis for a classification scheme, a grammar for automatically classing building maintenance, and lexicon for class headings. A sample of the vocabulary for classing job descriptions was discussed and possible difficulties identified. In the following Chapter the documents provided by Willmott Dixon Maintenance Ltd. are detailed and the effect of their sources upon the nature of the information they contain is analysed. Methods of transcribing job descriptions are also examined, as they bear upon computer processing, which is the subject of Chapter 9.

8.

Maintenance data.



Chapter 7 indicated how a classification of maintenance work and a grammar might be applied, and what the possible causes of noise would be. Examined here are first, the archived material provided by Willmott Dixon Maintenance Ltd. and its nature, then the broader issues associated with an unstructured source of building maintenance data, and finally, methods of their recording and transcription.

(a) Data population.

All the research material was provided by Willmott Dixon Maintenance Ltd., the collaborating organisation. A national company within the Willmott Dixon group, it provides a building maintenance service for every type of public, commercial, industrial and private client, except individual householders. Information was potentially available from any of eight regions, but the managing director, in deciding that with one it would be simpler to monitor the company's participation, selected the Midlands. This region is boundary-marked by Aberystwyth, Crewe, Mansfield, Aylesbury and Banbury, an area judged more than adequate to yield a characteristic sample of maintenance jobs. It covers approximately 10,200 square miles and includes several large conurbations equalling London in their diversity and density of buildings. A small team manages operations from an office and yard situated in Handsworth, approximately two miles north-west of Birmingham city centre.

(b) Sampling method.

Following discussions with the regional manager, a trial collection of jobs from current operations was carried out and revealed problems for both parties. Office staff, already fully occupied, had to interrupt familiar and practised routines to gather the information.

Having identified a suitable description, selected pages from the job's documents were photocopied and the papers returned to their place in the work-in-progress queue. The photocopies were then collated and confidential details censored. For the researcher, it was impossible to gauge the degree of exclusion, because only the selected job details were open to inspection.

(c) Confidentiality of records.

A wider examination of the invoicing process led to an archive of settled accounts. The company allowed unrestricted access to these records and released the most recent into the care of the present author. This arrangement eliminated disruption of the office-work and the risk (although still seen by the company's management as a valid one) of revealing sensitive and potentially valuable commercial material. On the single occasion when a manager needed to consult a job record, the papers were received at Handsworth on the following day. Physical security of the loaned material was equal to that used by the company itself; confidentiality of the information was maintained by restricting access to this researcher and on a 'need-to-know' basis for the Director of Studies and both Supervisors. This archive (except for a very small number of descriptions, whose origins are cited) was the source data for maintenance tasks throughout this enquiry.

(d) Identifying records.

The records were stored in natural arithmetic sequences of between 100 and 250 jobs, held together in paper cases. On receipt of an instruction, the repair was given a serial number, prefixed with letters identifying the issuing office, generally MB (Maintenance Birmingham). This brand identified every document to do with that job throughout the progress of its active life, onwards into invoicing, settlement of account and finally,

storage in its number sequence within the original series. No attempt had been made to sort, or classify the jobs in any other way, resulting in a randomness within limits imposed by whatever range of building types were owned, or professionally managed, by Willmott Dixon's clients. This is a highly significant feature of the data. Work was done in response to the needs of client organisations, some of whom could well have requisitioned repairs as planned or periodic maintenance. Others, evinced by the records, placed orders on an *ad hoc* basis, to meet commitments and obligations to tenants of their property. Thus, priorities of jobs were client-lead, on a continuum between those scheduled several weeks in advance, and emergencies demanding immediate attention. Intervals between arrivals of instructions were not timed, but are almost certainly random since they are impossible to control, giving rise to a queue. Contractors try to impose discipline upon the queue, with the intention of reducing customer waiting times as well as improving utilization of their resources. One strategy is to influence clients' perceptions of urgency, in this case, by having a sliding scale of call-out charges, a longer notice attracting a lower figure; working on premium time (outwith normal hours) markedly increasing labour costs. Other, more direct methods, are to increase the number of operatives (servers) up to a point just short of eliminating the queue altogether; operatives can keep their base updated with expected completion times, allowing them to be directed to jobs in their vicinity, cutting down travel times (favoured by Willmott Dixon). Given the records' age, although short, the effect, or extent to which these or other, feasible strategies were applied, could not be quantified.

Setting aside these conjectures and accepting that some measure of queue discipline was applied, and further, resulted in work-in-progress being perfectly balanced with demand, Willmott Dixon would *still* be unable to determine the *kind* of jobs entering the system.

The range of clients, the diversity of their building stocks in age, type and use, taken together ensure randomness in the presentation of maintenance tasks.

(e) Variability

Randomness is an attribute applicable to the intervals between jobs entering the work-in-progress queue, rather than the diversity of its content. Evidence for, or against the latter is slender, both of investigations and in quantitative terms. (Hague, 1983, p. 7.6) but what there is, indicates a recurrence of 'standard tasks', interspersed with the unusual (and therefore memorable) which bolster belief in the distinctiveness of every job. This conviction in the variability of maintenance work is not confined to building; the engineering industry, with generally more controllable working environments, has for many years considered it an established fact. Hurlston (1967) in presenting an advance in applying measured incentives techniques to this field, blamed the complexity of maintenance and indirect work for any lack of development during the previous twenty-five years, adding,

'A major difficulty is the quantity of non-repetitive or "one-off" jobs met with in maintenance work. Even when recurrent jobs of large work content do occur they are invariably spaced with long time intervals between them and if not supported by very precise information...they are also put into the class of "one-off" jobs.'  
(p.566)

Acceptance of maintenance being fundamentally variable, is, for the production engineer, an unchallengeable doctrine. Moreover, after a decade when several statistically based estimating methods were enthusiastically promoted, for example Kirby (1970), it is no longer considered a matter worthy of examination. The same faith in the heterogeneous nature of maintenance is commonly held by building practitioners, but without a remotely similar history of analysis, argument and publication. This present work does offer an observation; taking records in any order, it was never possible to predict the

next job description. That in itself, would satisfy any set of criteria for defining randomness, equal to the axioms cited by Ashford (1977);

'...rules concerning the relationships between the abstract mathematical concepts of probability and events are assumed to be true without proof.' (p. 87)

(f) The archived material.

Most of the job records were formatted as shown in Appendix D and the following details, stapled together and, as described (para. (d)), archived in sets of serial numbers

Copy Invoice

Section headings

To: (Address of contracted client, for invoicing)

From: (Willmott Dixon branch generating the invoice)

Invoice Number: Due for payment: Customer order: Contract: Customer:  
Invoice/Tax point: (Accounting control information. Job No. against 'Contract' is used in this present research to identify job description)

Unheaded section (Job details are typed here. Job address; work carried out; tradesmens' time and rate; materials and other charges)

Some invoices were typed, as evidenced by carbon copies with over-typings where corrections had been made. (See para. (g) below).

Daywork Sheet:

Section headings:-

Client. Date. Site Address

Complete By (Time limit for completion of job, i.e., it's priority level)

Job No. (Preprinted - see Input Copy)

Work Completed. (Description of repairs, action taken etc.)

Comments/Further Action. (Informs client and/or office of further problems, and need for other trades, materials, components)

Work Completed (YES/NO boxes to confirm whether or not job is complete)

Labour and Transport (Date. Names. Start and Finish Times. Hours.)

(Date. Vehicle Reg. No. Mileage. Hours Travel.)

Materials (Lists materials used, with costs if known)

Plant Hire/Skips (Lists plant and rubbish skips used)

The Daywork Sheet is signed by the operative and countersigned by the client or an authorised representative.

Confirmation Note (Raised in the office and completed by operative)

Section headings;

Client Name. Address. Date.

Description of Job (Completed by operative. Space for a fuller description than the Daywork Sheet)

Trade. Name. Time On. Time Off. Total Hrs.

General Comment (Supplements comments on Daywork Sheet)

Client Authorization (Signature. Name. Job Title. Completed by the client)

This form was printed on NCR paper in triplicate:- 'To Be Retained', 'To Be Attached To The Invoice', 'To Be Given To The Client'. Frequently all three were filed together. Confirmation Notes were only used in this research to furnish information not given on the Copy Invoice, Daywork Sheet, or when either, or both, of these two were missing.

Progress Copy (Used in the office for controlling and progressing the order

Never used as data source, because it contained costing and invoicing details.

## Input Copy (As for Progress Copy)

## Invoice Control Copy (As for Progress Copy)

This form, Input Copy and Progress Copy were padded together (in quadruplicate with the Daywork Sheet) and pre-printed with the Job Number, the only contribution they made to the data.

### (g) Inputting data by scanning.

The original intention was to collect the job descriptions by using electronic scanning and optical character reading (OCR) software, encouraged by the copy invoices in the first batch of records, which were typed on plain paper. After an extended trial, the method was changed to conventional keyboard input of the text and data storage on hard disk, with floppy disk backup, in case of loss or machine breakdowns. There were several reasons for this decision and they are detailed below.

Scanning with the only flat bed machine available, an Olivetti Pagescanner, was a slow process. Limiting the number of lines scanned to the text, left unwanted characters exposed (mostly charge rates and costs) and necessitated the use of a mask with an adjustable "letter box" perforation. This allowed only the job description to be read, but entailed a fresh adjustment for every change in the typescript; margins, length and number of lines etc. Typescript also had to be parallel with the scanning head, otherwise the OCR software would not recognise the letters. Further, since the required page was stapled to several other documents, one either had to fold the papers over at the top corner, arrange the mask before placing them face down on the glass plate of the Olivetti, at the

same time ensuring the wanted page was positioned correctly, or, separate the papers first, then collate and restaple them once an acceptable reading had been obtained. This latter method did make it easier to align the typescript and square it with the edge of the plate (the typing showed through the paper, if somewhat faintly), check the mask's position, close the cover lid and start the scanning head. However, this advantage could be lost if the work moved when closing the lid, necessary for keeping the paper in close contact with the glass scanning bed. With a sheaf of papers, the required overall pressure could be obtained with the flat of one hand, whilst operating the computer with the other.

The scanned descriptions were transferred to a 3.5" floppy disk, via copying from the 5.25" disk used in the Pagescanner machine configuration, and edited in a later operation. There were four types of fault to deal with before transcriptions could be considered acceptable; the OCR program produced double-spacing; overtyping where the typist had corrected an error on the top copy, presumably using a correction ribbon or fluid, but without removing or covering the carbon paper; missing characters and sections of text that could not be read due to poor quality copies; attempts by the OCR to interpret marks and blemishes arising from handling and storing the documents.

Operation and travelling times had to be extracted from the Daywork sheet, checked against the Confirmation Note and converted to minutes.

The nature and physical condition of the material was variable. Computerised, or semi-automated invoicing meant that tractor-fed NCR paper was used and copies



(sometimes on blue paper, well-known for the poor quality photocopies it produces) were printed diagonally across the space assigned for the description with the word COPY in 140 pt type. In some batches of job records, the typed copy invoices were missing, leaving handwritten descriptions on Daywork Sheets or Confirmation Notes, with, on occasion, expanded or edited versions, with all that entailed in poor legibility and spelling, absent punctuation and some novel syntax. Still other batches may have come from firms taken over by the collaborating organisation as going concerns. Ostensibly, their practices and methods had persisted, at least for a time, or resurfaced occasionally in their authors' phrasing, or spelling and use of particular terms and abbreviations.

(h) Manual data input.

The descriptions were typed on any IBM compatible desk-top computer using the Editor of Microsoft DOS 6.0. Each batch of jobs was allocated a subdirectory, named from Roman, Greek and Norse mythology (Bulfinch, 1963) simply because most meet DOS guidelines for file-names. They have eight characters, or less; will not contain illegal characters, such as semi-colon or '\', are distinctive and unlikely to be used in other databases. Every job was entered as a file in the subdirectory and identified with the Wilmott Dixon job number and the file extension, TXT.

Rules were formulated in order to create and maintain, as far as possible, consistently accurate and faithful transcriptions of the described work and integrity of their subsequent treatment.

They are as follows:-

type the format used on the copy invoice, thus, if upper case or normal typescript as presented. Overtyping, or typing a character in an obvious space between two

words, then use the correct term,

spelling errors, terminology, syntax etc., retain as presented,

handwritten descriptions, typed in lower case, with upper case characters as presented, except that blocked lettering was treated as if it were normal, discursive handwriting,

fractions converted to decimals,

several tasks under one job number only separated if times could be allocated,

times over 10 hours duration were described, at least in outline and the phrase "Job too large" followed after double line space with "Time out of range.",

craftsman times shown as "OT ..mins," assisting labour as "LM..mins," and the travel time as "TT ..mins."

Notes of misspelling, unusual terms or trade names, were made as they occurred and entered after every typing session in a file named in a way to link it to the subdirectory, and given the extension, .DAT.

Transcribing the documented descriptions in this manner was a slow and tiring process. It demanded concentration and care, with constant reviews to maintain, as closely as possible, adherence to the rules for transcription. There were, however, tangible advantages, which arose from the need for confidentiality. The task could not be delegated to anyone else, so this researcher read every record before, after, and during entry. Unfamiliar terms and abbreviations, the kind of language used, the idiosyncracies in style and phrase patterns, all were experienced at first hand, giving both breadth and depth to an understanding of their authors' vocabularies and patterns of English phrasing and grammar.

## Summary

The broad conclusions from this examination is that real world data are disordered, and subject to variation in the quality of the recorded information, but it is still possible to collect them in a consistent manner. In the next Chapter, a method for abstracting job descriptions from an established office computer system, then classifying them automatically, is suggested as a basis for further research and development.

9.

Processing examined.

The last Chapter described the mechanics of data collection and Chapter 7 presented a classification grammar for classing randomly generated descriptions of repair work. Here the two are brought together and experiments in programming to test the feasibility of applying them are described and discussed. Finally, a method for automatically abstracting and classing data is suggested, and its functional requirements are defined.

(a) Computer programs.

Programming is not part of this present research, and the following discussion is intended only as a statement about software for the suggested system, or future avenues of enquiry. Two computer languages are proposed: QBASIC for manipulating the original job descriptions, and, Prolog for translating sentences and organising their classification. but others might prove equally effective. Program modules were written, the majority in QBASIC, for experimenting with descriptions taken from the transcribed samples (Chapter 8). Successful outcomes would, by extension, indicate support for the research hypothesis.

QBASIC is a development by Microsoft Corporation of Quick Basic, which itself is a variant of BASIC, an acronym for Beginners All-purpose Symbolic Instruction Code. Written in 1964 by Professors Kemeny and Kurtz, Dartmouth College (Forsyth, 1978, p.1) as a method of introducing students to computer programming, it has become available on most personal computers and more machines are sold with BASIC installed than any other procedural language (Perry, 1993, pp. 10-13). Current versions have all the advantages of a structured code, whilst retaining the directness of the original. Specialist advice was sought from academic users about possible alternatives: opinion held that any modern variant would be suitable for the envisaged processes. Indeed,

James (1984) demonstrated in his practical work on artificial intelligence (AI):

'...using a language such as BASIC does at least prove that AI techniques are general programming methods and don't depend on any special facilities available in languages, such as Lisp and PROLOG that are more often used in AI.' (p. 8)

Prolog is a rule-based language, developed for programming logic, but although the first versions appeared in the early 1970s, demands made by the code meant that personal computers did not have sufficient processing power, or dynamic memory capacity to execute programs at an acceptable speed, until 10 years later. Prolog, like BASIC, has many variants, but one implementation, Edinburgh Prolog, is described by Clocksin and Mellish (1987) as a '*de facto* standard' and is the version used to run the demonstration programs (Appendix E, pp. L - LIII) for classing job descriptions.

(b) Experiments with programs.

The experiments for the proposed classification system was conducted , as described above, by writing short programs in both languages. No attempt was made to model a complete application, just to test whether the adopted computer languages could meet certain of its perceived functions. Some could not be assayed. For example, the ability to switch from one processing language to another, which must be a feature of the 'unseen' classifier, was impossible to recreate at this stage. Because sentences could only be processed, by either standard personal computers that run only one program at a time, or a more powerful machine capable of running two programs at once, but required the development of what, in effect, would be the working application. This situation was a consequence of the present author having limited access to a computing science laboratory with the technology and, more importantly, the expertise to give advice on this specialised area of computer programming. Fortunately, the same facility offered the

opinion that this particular need for parallel processing was within the bounds of current practice and knowledge.

Each of the experiments next described, sought answers to a limited range of questions, not necessarily in the sequence of a fully developed program. Input was generally via the keyboard, and all output was displayed on a terminal console or monitor of a personal computer. This method gave results that could be read immediately. When an error was detected, the program terminated, and the fault highlighted on the screen for the necessary corrections to be made, an iteration allowing many trials in the course of relatively short sessions. There is a further advantage; writing, or typing to a screen is, in computing terminology, 'printing to a device'. And since printers, fax machines and disks for storing data, are all 'devices', then one item of hardware is equally valid as a means for receiving the form of output intended to be hidden in the auto-classifier.

The first program listings (Appendix E, pp.XLI - XLII) are responses to seemingly pedantic impositions of Prolog's presentational requirements. Clocksin and Mellish (1987) counter this perception in the opening pages of their text, emphasising;

'...the following things are important:

The names of all relationships and objects must begin with a lower case letter. For example, likes, john, mary.

The relationship is written first, and the objects are written separated by commas, and the objects enclosed by a pair of round brackets.

The full stop character "." must come at end of a fact.' (p. 3)

The initial question posed was "Could QBASIC change natural sentences and satisfy this format?"

(i) Experiment 1

Converting text to lower case is a simple matter of using a QBASIC function, but removing unwanted punctuations, inserting commas, the end period and brackets is more involved. QBASIC will delete a character from a sentence (word-string), but leaves a space (treated as a character). Conversely, a space can be found and a comma inserted. If there are two adjacent spaces, the result is two adjacent commas. The program on page XLI changes a sentence input at the screen prompt;

**Type description—ATTEND SITE, CLEAR AND ROD DRIANS. RUNS OK**

**result = attend site clear and rod drians runs ok**

An idiosyncrasy of the program, removed the comma and shortened the word-string, but exchanged a space for the period after 'drians'. Another QBASIC function can match characters in a word-string, returning the number of its position. Double spaces can thus be identified and one of the spaces removed, an action performed for the first occurrence only. Succeeding matches have to be made by inputting the sentence again, in a new, shortened, form. However complicated this might appear, programming is relatively straightforward, allowing the sentence about 'DRIANS, to be processed with output;

**result—[attend,site,clear,and,rod,drians,runs,ok]**

and is very close to Prolog requirements.

(ii) Experiment 2.

Job descriptions had been transcribed and kept closely to their original form as far as possible, in batches of up to 200 files, and stored in directories. (Chapter 8) A similar situation might arise in a developed system, if descriptions were abstracted for processing at a later stage, rather than singly as envisaged. Prolog, like any computer language, has to be instructed where data is stored, and this experiment tested whether



QBASIC could collect a description, and by matching keywords, identify the class or category heading (as far as Level 1), then write the data to another program file associated with the relevant lexicon, before directing Prolog to complete the classification.

The program DATEST2.PRO (Appendix E, pp. XLIII - XLV) reads files, one at a time, until the end of the designated directory, matches key words and totals a value JU that is used to calculate a probability function and denote a class heading.

```
C:\DOUGLAS\COMPONEM\5631.TXT is the file being 'presented' for your inspection.
Note that keywords, when identified, are shown at position P and
printed before the line of text in which they have been found.
Although this is not invariably so after amendments on 24/5/96

*****

to attend site and take details of materials required,
collect from suppliers and return to site.
replace door handles to outside door and put back
:- Word door at 9 JU= 2 Total 2
Line value of JU = 2
:- Word handle at 14 JU= 37 Total 39
Line value of JU = 19.5
handle to inner door. temporary used on outer door.
:- Word door at 17 JU= 2 Total 41
Line value of JU = 13.66667
:- Word handle at 1 JU= 37 Total 78
Line value of JU = 19.5

ot 45mins, tt 30mins.

Average of JU 45 and
```

Figure 9.1 Screen output -Job No. 5631

Thus, Job No. 5631 was transcribed;

To attend site and take details of materials required,  
collect from suppliers and return to site.  
Replace door handles to outside door and put back  
handle to inner door. Temporary used on outer door.

OT 45mins, TT 30mins.

and as Fig. 9.1 shows, the terms 'door' (JU = 2) and 'handle' (JU = 37) were each identified twice, but on different word-strings

Similarly, Job 5690 was transcribed;

To attend site and take down curtains and track  
and tract fixing batten.

replace 1 x 1 batten with 2 x 1 and screw fix securely  
to wall. REfix track and rehang curtains.

OT 120mins, TT 105mins.

and displayed on the screen in Figure 9.2. Notice 'track' (JU = 35) matched twice, 'wall' (JU = 4) once. In a full version, the 'Average of JU 45 and' appearing at the bottom of the screen, would be a correct value and used to calculate a probability function of the class of work, which is demonstrated in the next experiment.

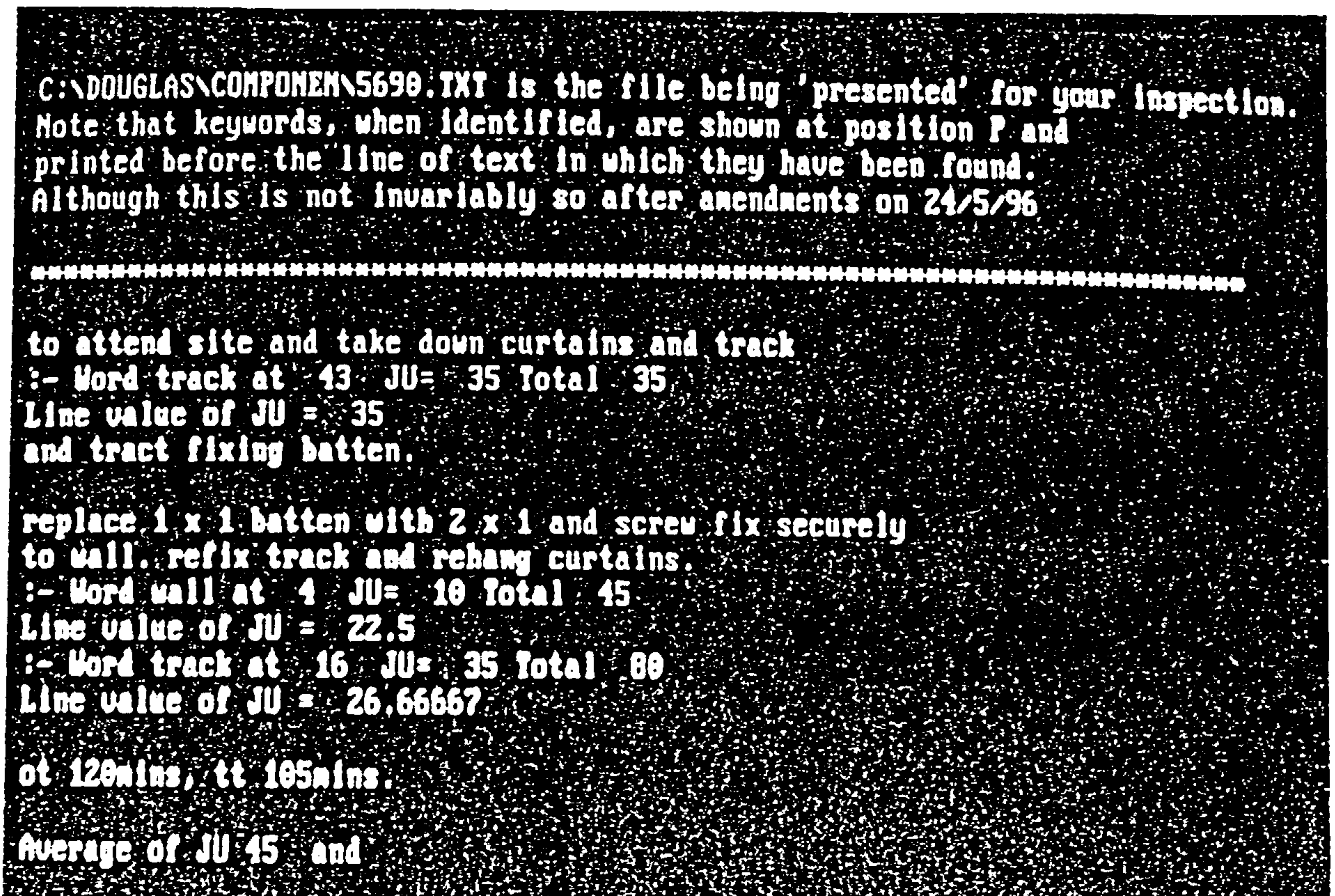


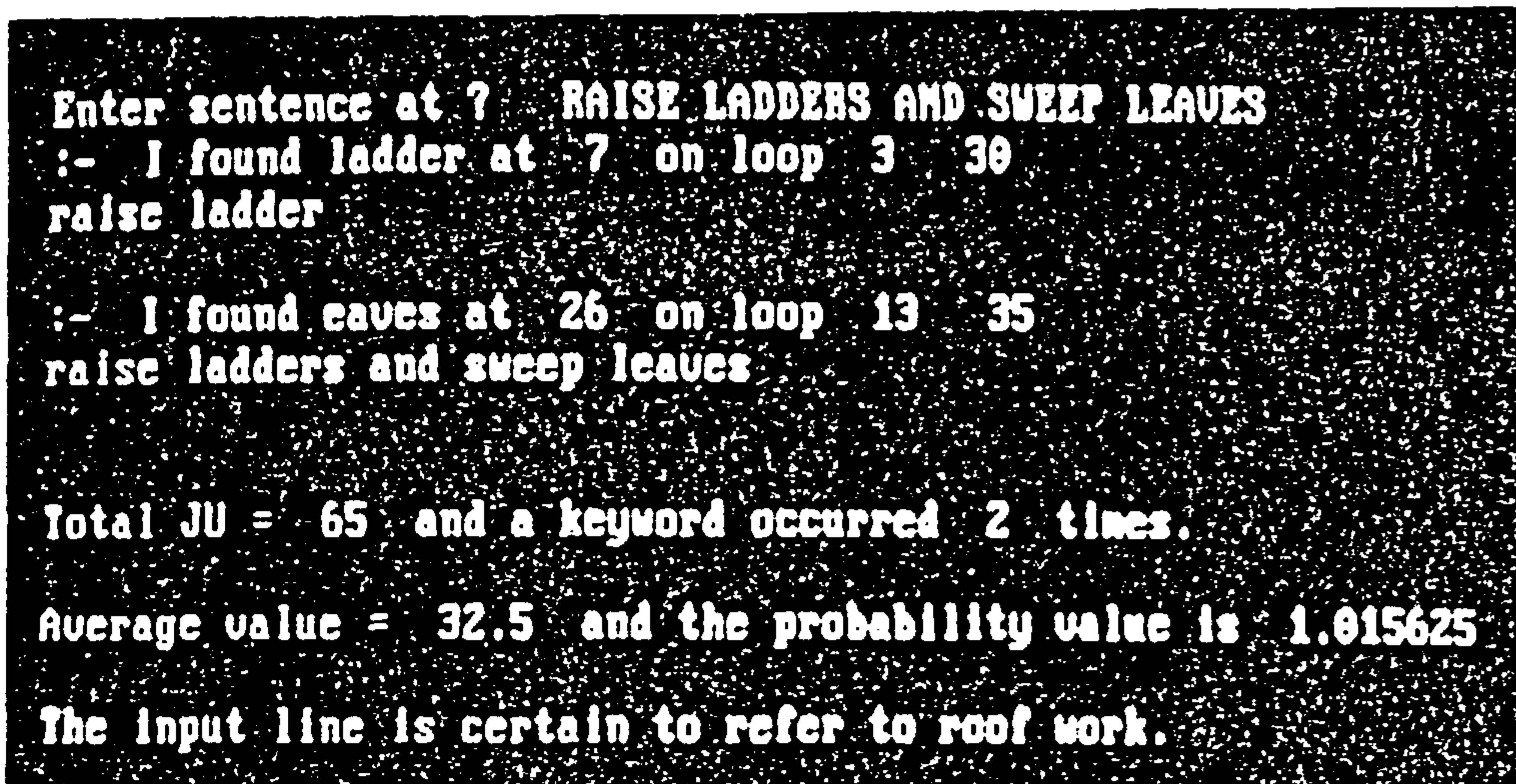
Figure 9.2 Screen output- Job No. 5690::

(iii) Experiment 3.

The program D2TATEST.PRO (Appendix E, p. XLIX) reads a sentence typed to the screen via the keyboard, calculates a probability function for the class heading and prints a sentence giving a degree of certainty. The first input, entirely in upper case was:

RAISE LADDERS AND SWEEP LEAVES

QBASIC matched 'eaves' in 'leaves' giving a false, but fortuitous, certainty that the work must be to a roof. The important lines of output are, 'Average value = 32.5 and the probability value is 1.015625' 'The input line is certain to refer to roofwork' 'Average value' is the equivalent of JU in Figures 9.1 and 9.2 and the 'degree of certainty' is intended to signal to Prolog that the description will be found under the heading ROOF in the classification.



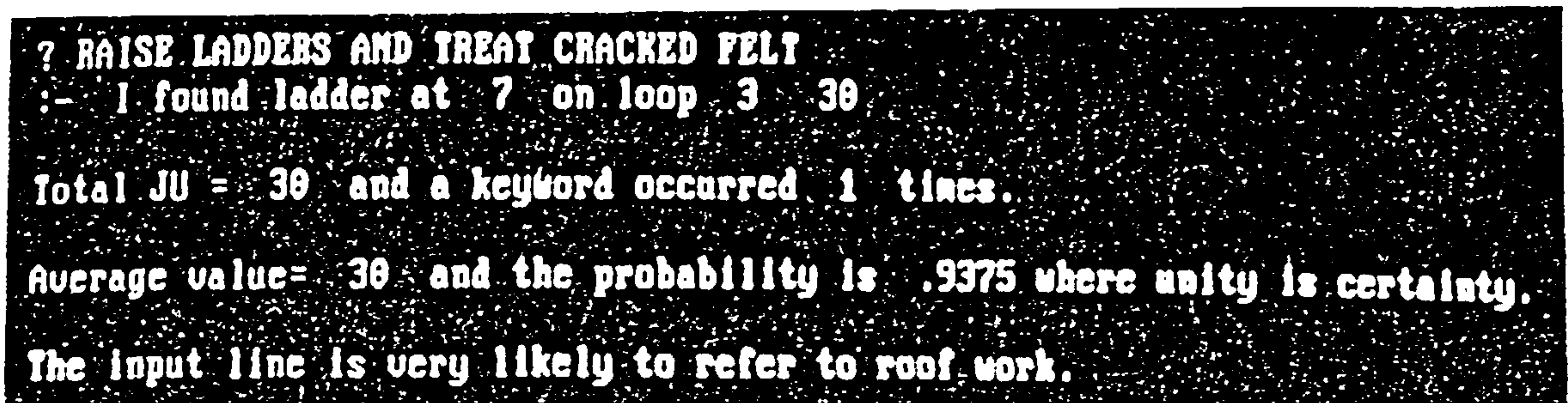
```
Enter sentence at ? RAISE LADDERS AND SWEEP LEAVES
:- I found ladder at 7 on loop 3 30
raise ladder
:- I found eaves at 26 on loop 13 35
raise ladders and sweep leaves
Total JU = 65 and a keyword occurred 2 times.
Average value = 32.5 and the probability value is 1.015625
The input line is certain to refer to roof work.
```

Figure 9.3 Example testing probability -first case.

A similar, but slightly less definite outcome, results from typing the line;

RAISE LADDERS AND TREAT CRACKED FELT

due to the keyword 'ladders' being the only term recognised. Adding further preferred terms, and their synonyms, to the DATA statements, also revising IF THEN tests for 'true' and 'false' conditions (p. XLIX), will improve the separation between classes.



```
? RAISE LADDERS AND TREAT CRACKED FELT
:- I found ladder at 7 on loop 3 30

Total JU = 30 and a keyword occurred 1 times.
Average value = 30 and the probability is .9375 where unity is certainty.
The input line is very likely to refer to roof work.
```

Figure 9.4 Example testing probability -second case.

(iv) Prolog output

Output from the programs written in Prolog are not depicted, due to the unsuitable conditions for taking photographs of screen displays in the computing science laboratory.

A full listing of a demonstration program is included in Appendix E (pp. L - LIII), and as the commentary, inserted by the program's author, Dr P. R. Innocent explains,

'The program is not commented as a working knowledge of prolog is necessary to understand fully the semantics of the program.' (p. LI)

supporting the reason for not discussing the results, beyond pointing to the input sentence;

'To attend site and rod all manholes on drain run to clear blocked WC.' (p. L)  
and the result 'YES'.

(c) Criteria for an auto-classifier.

A framework of the processing system was drafted before seriously attempting to test the hypothesis that reducing job descriptions would lead to their automatic classification.

Anticipating discovery of supporting evidence allowed, an outline for a practical application to be defined, giving direction and added substance to the linguistic objectives. The attributes judged to be desirable were that::

the system must function independently without human intervention, except when a report is called for,

inputting data must not place additional burdens upon a user, beyond setting parameters for the output, or directing a classification of particular data files,

the system will run on standard computing hardware, possibly a local area network (LAN) as a maximum,

it must be a software solution, kept within the bounds of current computing science techniques and knowledge,

commercially available software is used where possible..

Assembling the data for this present research was a discrete operation (see Chapter 8), but in an operational context, descriptions might be processed in batches, or preferably, initiated when they enter an invoicing procedure. For example, a suite of programs placed in the accounting system, could analyse the information and store the results until a report was requested. Changing from this primary research to the subordinated role in an organisation would entail minor changes to that part of the program handling data input, and tailoring the reporting and formatting to suit specific administrative procedures.

(d) Concept of a developed system.

Normally, the translation of sentences is carried out more directly than in this present research, and Fig. 9.5 shows the difference. Taking a sentence  $S_1$ , it is reduced and

transformed into a simplified form and matched with terms in the 'heading lexicon', to establish a class heading,  $S_{Head}$  at least to a category at Level 1. (see Chapter 7 (e)). Then,  $S_1$  together with  $S_{Head}$  are processed again. If the rules (conditions defining a class or category) confirm instantiation, the job is correctly classed.  $S_2$  the sentence denoting the repair is written by an 'Inference transformation'. The annotation, 'Normal/direct route' is the usual sequence for parsing and transforming  $S_1$  into  $S_2$ , but without establishing a context (job class). Fully developing the 'Inference Transformation' rules that will enable inferences to be drawn from job descriptions, possibly written in several sentences, and ensuring that a job description can be classed correctly, is outside the bound of this research.

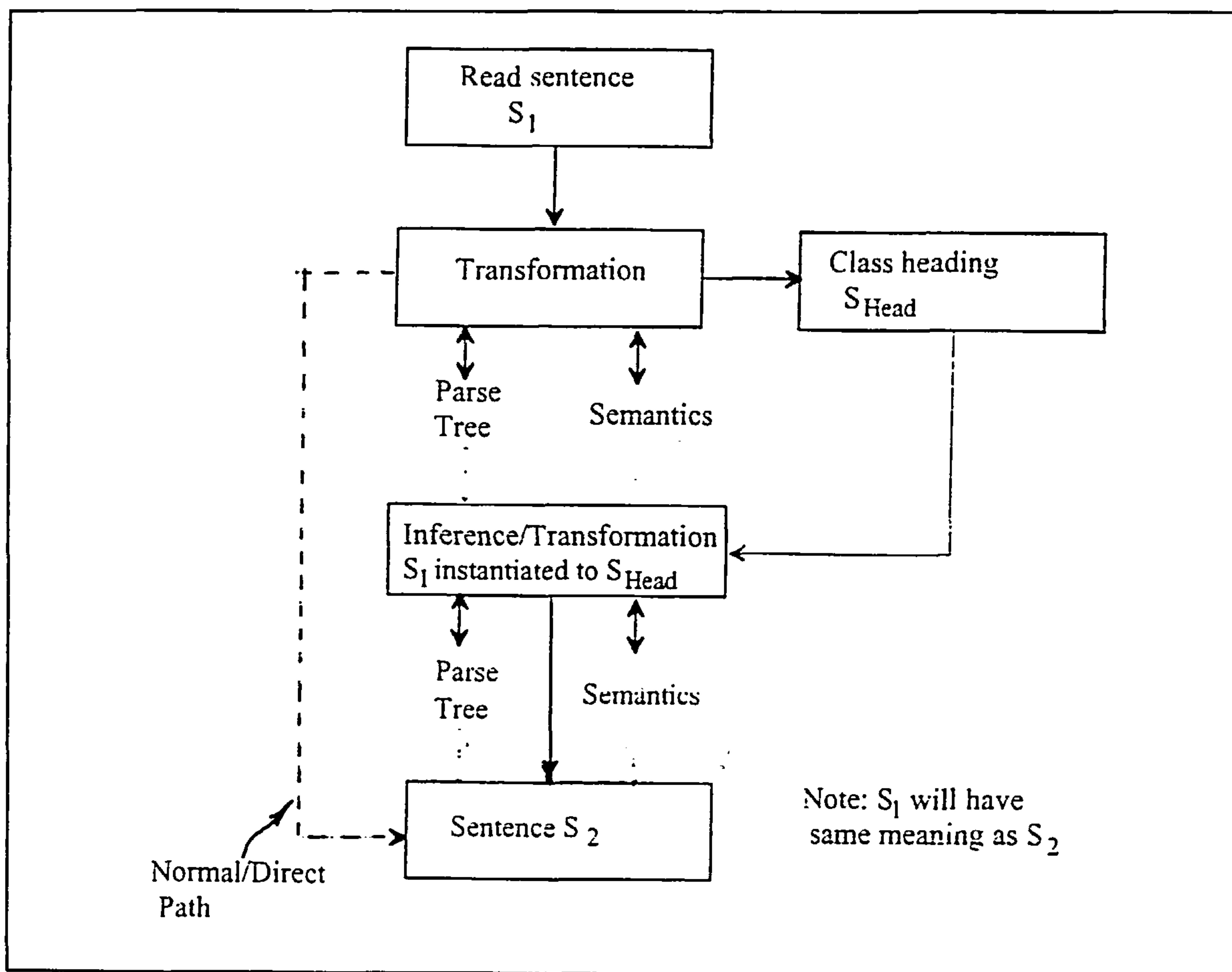


Figure 9.5 Comparison between normal approach to transformation and suggested method of auto-classification

(e) Organisation of an auto-classifier.

Computer processing of natural languages is a complex problem, and the difficulties presented by a limited sub-set of English have been detailed in earlier chapters. The organisation suggested for a system, is depicted at Figure 9.6, in a flow chart to describe the classification process of a single job description.

(i) Read invoice description. Reading the description from a 'data-base' depends upon where the system is positioned. The words can be copied from the invoice when it is typed into the costing system, or read from a batch file of completed invoices.

(Transcribing from paper copies, described in Chapter 8, is not envisaged as a practical method).

(ii) Description for headings. This operation can be controlled by QBasic. The invoice description is converted to lower case, necessary to satisfy Prolog input. Words are compared with terms in the 'Lexicon of headings', if matched, a class heading is built up with the sentence structures (S→N; S→A+N;...) discussed in Chapter 7 (e). In a few instances during the trials, for example when words happened to match a relevant ordered sequence of terms, QBasic built headings to Level 2.

An automatic spelling check was considered that would identify and correct misspelt words by matching phonemes to syllables (*roove* = *roof*). Because the available word-processing software only lists possible alternatives, an operator's decision is required, suspending the invoicing procedure. The complications inherent in this are sufficient to call for further research.

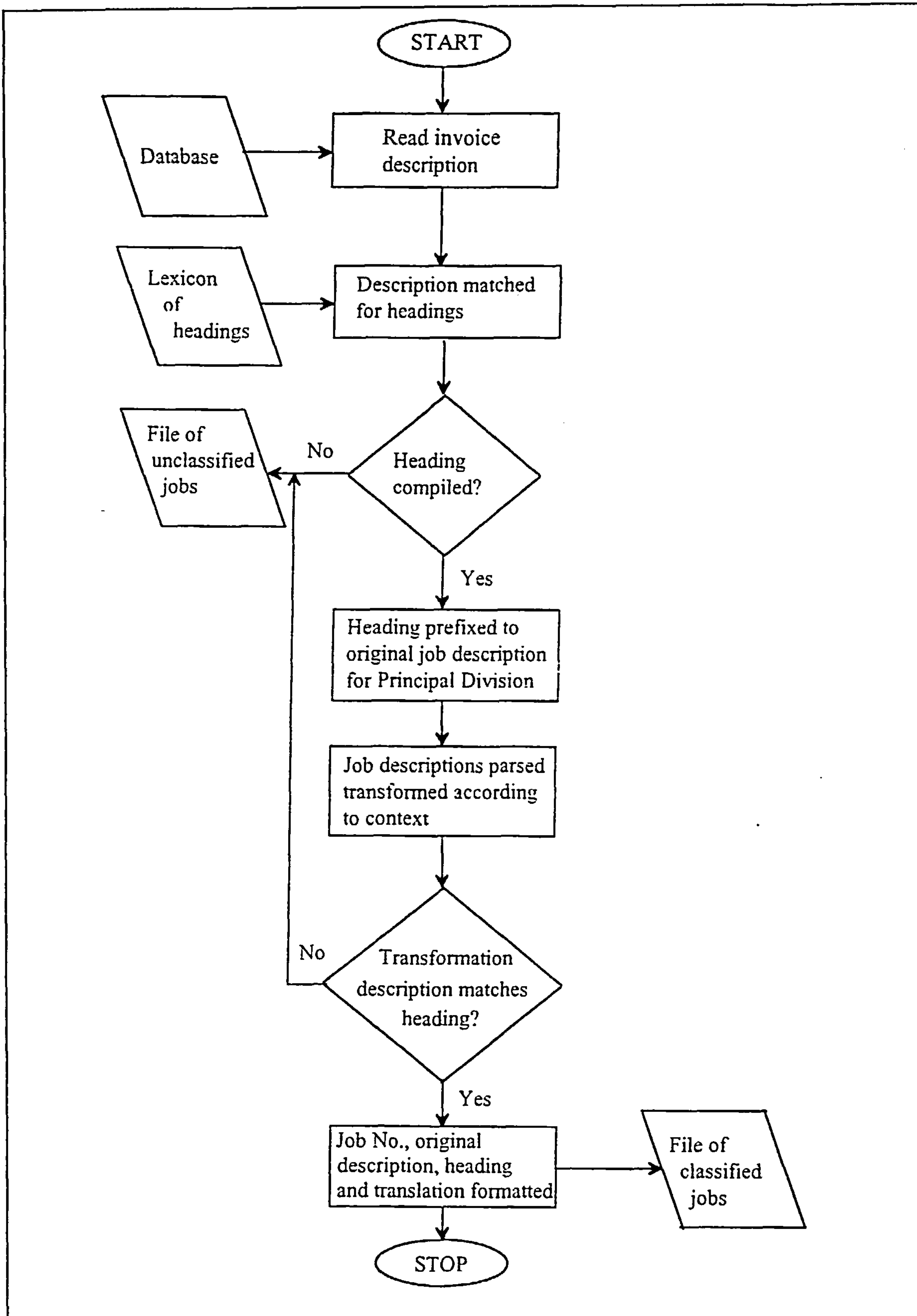


Figure 9.6 Flowchart of automatic classification process.



(iii) Heading compiled? There are two possible results:

'No-heading' and the description passes to the 'File of unclassified jobs',

'Yes-heading' and the description and heading move to the next operation.

(iv) Heading prefixed.. An operation under QBasic control. The original description (still in lower case) is prefixed by a Principal Division heading, identified with its job number and details of the operation times added. At this stage the description has been placed in a context for the translation process.

(v) Job description parsed... Control passes to Prolog. The description is parsed and transformed, a vocabulary for the applicable Principal Division being incorporated into the Prolog program. Transformation follows the grammar based upon the small number of verb-phrases discussed in previous chapters.

(vi) Transformation..matches heading? Whilst the contextual vocabulary is intended to be specific, it will also have some terms in common, and a well-formed sentence might be parsed by a chance congruence of terms. This is another *a priori* decision and a non-match adds the description to the file of unclassified jobs.

(vii) Job, original description... formatted. A routine, possibly under QBasic control, formats the output according to the parameters set by the responsible manager. There are many variations and features that can be incorporated at this stage, including statistical monitoring, financial controls, incentive payments, and to bring this proposal back to the reason for this present research, matching the repairs against scheduled rates.

It must again be emphasised that this is not an attempt to present a definitive solution, merely to indicate the one possibility set in relationship to the data and the linguistic concepts found within them.

### Summary.

This Chapter has described the experiments inputting data from the sample of transcribed job descriptions, testing the feasibility of the adopted computing languages, for classing, exploring some of their limitations and strengths. The framework for a system capable of classing descriptions automatically has been anticipated and its functional requirements defined. Further research seen to be necessary before a marketable application can be developed.

10.

Conclusions and recommendations.

(a) Conclusion: the hypothesis

Collecting information from building operations is not a practical matter, easily resolved by adopting what Kelly (1826) entreated, a 'rational mode of delivering accounts', it is a complex theoretical problem, inherent in the nature of construction work. Confining the field of research to small-scale maintenance work has, paradoxically, increased the order of difficulty. Large-scale building works can restrict intelligence gathering to trained staff, and the approach in estate management is to use schedules of standardised descriptions, a solution because contractual arrangements impose their use. But for the majority, the building firms providing a service for a variety of clients, reports written by operatives are the vital source of feedback for management decisions and administrative operations; the quality of such data is unreliable, and their total volume can overwhelm any attempt of rational analysis.

This research sought to improve site feedback, by answering one question: Can data be abstracted from the workplace, automatically, by the simple expedient of identifying a limited number of words in descriptions of operations?

That computing technology is capable of translating written text from one language to another is undeniable, but depends upon having an established, properly constructed, grammatical source. Invoices and work records were not always written in 'good' English, and their translation depended upon the reader's knowledge of craft processes and using one sentence to interpret the next, a facility not given to machines.

Analysis of descriptions, transcribed from Willmott Dixon Maintenance Ltd.'s invoice records, revealed several linguistic patterns, the most significant being a verb-phrase,

verb + adjective + noun. expressed in such sentences as, "Repair damaged post", "Replace roof tiles", and became the principal indication supporting the hypothesis. Applying the principles of a Minimum Vocabulary, gave further impetus for reducing grammar to a primitive level, <sup>OGI</sup>helped by accepting that the human reader could infer meaning from words, whilst a computer can only match signals denoted in 'word-strings'. Although the problem had been considerably simplified, there are many linguistic contradictions remaining to be resolved, before a complete set of theoretical principles for translating job descriptions might be established.

(b) Conclusion: a classification.

Deleting prepositions, conjunctions and proper nouns from the proposed translation grammar, lead to a classification that itself was a grammar, with two syntaxes. The first, and simplest, used only nouns, which programs written in QBASIC, were able to manipulate to a degree that pointed to a productive outcome. The second syntax, supports the bulk of the classification, and contains most of the linguistic ambiguities and contradictions, particularly those posed by polysemic words, but the small number of verb-phrase patterns discovered in a large sample of disordered real-world data, together with the limited ranges of preferred terms, indicate a possible solution.

Published classifications of building maintenance were shown to be expressions of the organisation of local and national authorities, and this research is a reaction to their use of standardised descriptions and codes. Nevertheless, given these databases of standard descriptions, there is sufficient in this investigation to show that identifying descriptions with a limited vocabulary is possible.

(c) Conclusion: automatic classification.

Although the concept of an automatic classification has driven this research, a full analysis of computing science theory and its application has been restricted to defining the parameters of a system, and was largely achieved in the previous Chapter. Whether an application could be developed as a direct outcome of this research depends entirely upon resolving such problems as misspellings, gerunds, transfer of meanings between sentences (Take down signs, found beyond repair. Remove) and polysemy. The results of experiments with programs, (Chapter 9) were sufficiently encouraging to indicate support for the lines of enquiry and research discussed in the next section.

(d) Recommendations.

The implications for this research are mainly practical and might apply to other sectors of construction outside of building maintenance. An example of generality is the large construction project: progress and cost controls are often based upon critical path programmes, activity descriptions forming the class headings for allocating labour and plant, materials and subcontractors. With a classification grammar and a vocabulary pertaining to the project type, reports written by supervisors could be directly transcribed, for the system to produce reports.

Provided there is a coherent body of data and defined rules for classing, then an automatic procedure seems to be feasible.

Recommendations for developing an application are shown in Figure 10.1. Each should be understood as discrete developments and ultimately, computer programs and databases that would be operated by a master program, or gatekeeper, controlling input and output, switching data from one section to another, and, crucially, ensuring the user

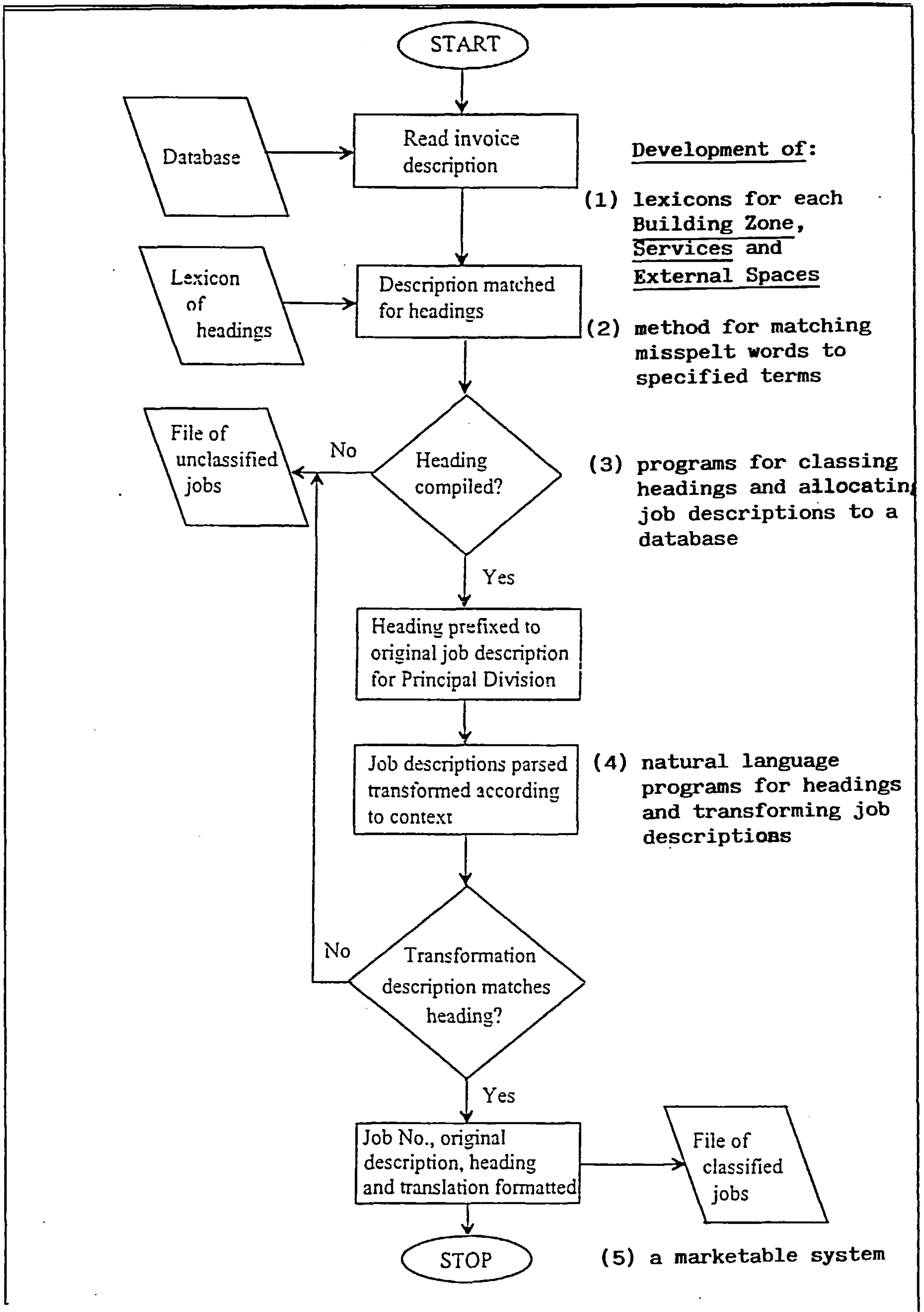


Figure 10.1 Research needs for the development of an automatic classification system

is denied access to the programs except under predetermined conditions.

1. Lexicons are the vocabularies of preferred terms, synonyms and trade terms. their compilation needing an extensive knowledge of building maintenance and the language used by originators of job descriptions.
2. 'Misspelt words...' is a continuation of lexical development, the problem being how to anticipate misspelt words not previously entered in the system. One avenue of enquiry would seem, from the evidence of this present research, is to use phonetics (vierbul = viable). Every word in the English language is potentially a candidate, therefore any solution must be general.
3. This is principally a matter for Computing Science. A procedural language (QBASIC) was envisaged as being the most suitable for preparing raw input for Prolog, extended to include the identification of class headings. Before making any decision about the continuing role of QBASIC in a full development, it is suggested that further research is undertaken to establish the limits of the language (or any of the variants).
4. All program development at this level is recommended to be a cooperative project between Computing Science and the knowledge and experience of practitioners in building maintenance management.
5. A marketable application is anticipated as a two stage development. First, running programs under manual control, correcting coding errors and editing preferred terms, an iterative process that should continue until the percentage of 'unclassified jobs' tends to zero. Second, market research is advocated, referring in particular to a general application in other sectors of construction.



## Summary

This research has addressed problems arising from written communications in building maintenance. Syntactic and semantic ambiguities have been partly resolved, indicating avenues for establishing a body of applied theory. Recommendations are made for the structure of an automatic classification system which, it is concluded, has support for further research and development.

# APPENDIX A

This Appendix contains extracts from the *Thesaurus for the Construction Industry*, First Draft - November 1970.

Compiled by Michael Roberts, Chris Eve, Peter Linn and Ellen MacHale

for the North Western Polytechnic School of Librarianship and the Brixton School of Building. (Now South Bank University).

(The final published version is understood to be Crown copyright, but was not referred to, or consulted at any stage by this present researcher).

NOTE:

The terms in the following schedule represent OPERATIONS, i.e. actions carried out upon materials, parts, etc. during construction by agents. They usually imply human involvement at some stage.

PROCESSES are actions which occur within materials, parts, etc. as a result of internal dynamism. Again agents can be discerned, but usually without human involvement.

Many terms representing Operations also represent processes e.g.

COLLAPSING of formwork e.g. due to overloading

COLLAPSING of telescopic formwork prior to movement  
to a new site.

POWDERING of mortar e.g. due to chemical reaction

POWDERING of aggregate in a pulveriser

It will frequently be found useful to collocate Operations and Processes represented by the same term.

Therefore the Operations facet has been extended in coverage so that the same terms may be used, either as Processes or as Operations.

Should it be found necessary to distinguish a Process as such, the Operations facet as a whole should be repeated under the class heading PROCESSES.

G0301	Operations	
G0311	Managing = Administrating	
G0313	Organising	
G0315	Co-ordinating	
G0325	Supervising = Superintending	
G0327	Permitting = Allowing	
G0329	Enforcing	
G0331	Exempting	
G0341	Theorising	
G0343	Investigating = Studying = Researching = Inquiring = Inspecting = Examining	
G0345	Surveying (i.e. making an extensive study of)	
G0347	Experimenting	
G0349	Analysing	
G0351	Synthesising	
G0361	Observing	
G0363	Identifying	
G0365	Comparing	
G0375	Interpreting	
G0377	Defining = Determining = Delineating	
G0379	Specifying	
G0381	Setting out = Marking out	
G0391	Measuring	
G0393	Batching (i.e. measuring of quantities) (Measuring by what measured)	
G0395	Weight batching	"
G0397	Volume batching	"
G0407	Calculating = Computing = Reckoning	
G0409	Adding = Summing	(Calculating by mathematical process)
G0411	Subtracting	"
G0413	Multiplying	"
G0415	Dividing	"
G0425	Squaring = Squaring up (i.e. areas)	(Calculating by what calculated)
G0427	Quantifying (i.e. quantities)	"
G0429	Taking off	"
G0431	Working up	"
G0441	Evaluating = Valuing	"
G0443	Costing	"
G0445	Accounting	"

(Operations)

G0455	Estimating = Assessing	
G0457	Predicting = Forecasting	
G0459	Testing	
G0461	Proving	
G0463	Verifying = Checking	
G0473	Planning	
G0483	Designing	
G0485	Architectural designing = Architecture	(Designing by what designed)
G0487	Engineering	"
G0489	Civil engineering	"
G0499	Structural engineering	"
G0501	Detailing = Detail design	(Designing by extent of design)
G0511	Describing	
G0513	Recording	
G0515	Drawing	
G0525	Communicating	
G0527	Instructing	
G0529	Programming (i.e. computer instructing)	
G0531	Conferring = Discussing	
G0533	Interviewing	
G0543	Advertising	
G0545	Indicating	
G0547	Warning	
	(Operations involving participation)	
G0647	Attracting	(Operations by nature of participation)
G0649	Repelling	"
G0651	Interacting	"
G0653	Coacting	"
G0655	Reacting	"
G0657	Counteracting	"
G0659	Resisting	"
G0661	Preventing	"
G0671	Protecting	"
G0673	Insulating = Shielding = Sheltering (i.e. cutting off from)	"
G0675	Lagging	"
G0685	Proofing (i.e. making impervious to)	"

(Operations)

(Operations involving participation)

G0687	Preserving	
G0697	Controlling = Governing	
G0699	Restraining (i.e. movement/action)	(Controlling by what controlled)
G0701	Constraining = Confining = Restricting = Limiting	"
G0703	Steering = Guiding (i.e. direction)	"
G0705	Operating	"
G0707	Mechanising	"
G0709	Automating	"
G0809	Changing = Altering = Varying	
G0811	Modifying	(Changing by degree of change)
G0813	Adapting	"
G0815	Converting	"
G0817	Fluctuating	(Changing by constancy of change)
G0827	Increasing	(Changing by change effected)
G0829	Enlarging	"
G0831	Expanding	"
G0833	Inflating	(Expanding by method)
G0843	Developing	
G0853	Decreasing = Reducing = Declining = Fading = Diminishing	(Changing by change effected)
G0855	Contracting = Shrinking	"
G0857	Lengthening = Elongating	"
G0859	Shortening	"
G0861	Standardising = Normalising = Equalising	"
G0863	Rationalising	"
G0873	Simplifying	"
G0875	Clarifying	"
G0877	Improving	"
G0879	Restoring = Repairing = Making good = Mending = Overhauling = Reconditioning	"
G0889	Deteriorating	"
G0891	Decaying = Decomposing = Rotting = Dilapidating	"
G0893	Damaging	(Deteriorating by cause)
G0895	Wearing	"
G0905	Preparing	(Changing by change effected)
G0907	Completing = Finishing	"
G0917	Starting = Initiating	"
G0919	Exciting	"
G0921	Stopping = Arresting = Halting	"
G0931	Accelerating = Speeding = Hastening	(Changing by change effected)
G0933	Slowing = Decelerating = Retarding = Braking	"

(Operations)

(Changing = Altering = Varying)

		(Changing by change effected)
G0943	Opening	"
* G0945	Closing	"
G0947	Sealing	"
G0949	Blocking = Plugging = Stopping	"
G0951	Caulking	"
G0961	Strengthening	"
G0963	Reinforcing	(Strengthening by method)
G0965	Bracing	"
G0967	Weakening	"
G0977	Toughening	"
G0979	Stiffening	"
G0989	Maturing	(Changing by change effected)
G0991	Curing	"
G0993	Ageing	"
G1003	Heating	"
G1005	Cooling	"
G1007	Refrigerating	"

(Changing).

	(Changing physical state)	(Changing by change effected)
G1107	Solidifying = Hardening = Setting	"
G1109	Freezing (i.e. liquid to solid)	"
G1119	Fluidising	"
G1121	Liquefying	"
G1123	Melting = Thawing = Fusing * (i.e. Solid to liquid)	"
G1125	Condensing (i.e. vapour to liquid)	"
G1135	Vapourising = Volatilising	"
G1137	Evaporating (i.e. Liquid to vapour)	"
G1139	Boiling	"
G1141	Sublimating (i.e. Solid to vapour)"	"
G1151	Gasifying	"
G1161	Thickening	"
G1163	Gelling = Coagulating = Fattening = Setting up	"
G1165	Feeding = Livering	"
G1175	Thinning	"
G1177	Diluting	"
G1187	Reacting (i.e. changing chemical composition)	"
G1189	Hydrolising = Hydrating = Slaking	"
G1191	Dehydrating	"
G1201	Oxidising	"
G1203	Burning = Combusting	"
G1205	Igniting	"
G1305	Surfacing = Dressing = Tooling (i.e. Changing surface)	(Changing by change effected)
G1307	Levelling	"
G1309	Smoothing	"
G1311	Polishing = Buffing = Burnishing = Lapping	(Smoothing by degree)
G1313	Rubbing down = Flattening = Sanding (i.e. smoothing by abrading)	(Smoothing by secondary operation)
G1323	Roughening = Hacking = Keying = Scoring	



(Operations)

(Changing)

(Surfacing)

(Changing by change effected)

G1333	Stone dressing	(Surfacing by object)
G1335	Hammer dressing = Bull facing = Clouring	(Stone dressing by tool used)
G1337	Picking = Pecking = Sparrow- pecking = Stugging	
G1339	Dabbing = Dabbling	
G1341	Batting = Broad tooling	
G1343	Boasting = Drowing	
G1345	Combing = Dragging	
G1355	Punching	(Stone dressing by resultant finish)
G1357	Scappling	
G1359	Nidging = Nigging	
G1361	Angle tooling = Angle drowing	
G1363	Scabbling	
G1365	Rusticating = Channelling	
G1367	Reticulating	
G1369	Vermiculating	
G1371	Reeding	
G1373	Fluting	
G1375	Furrowing	
G1377	Broaching	
G1379	Fitching = Rock facing	
G1479	Sharpening	(Surfacing by object)
G1481	Honing	

(Operations)

(Operations involving no change)

G1581	Holdings = Retaining	
G1583	Gripping	(Holding by force used)
G1585	Clamping = Cramping (i.e. holding together)	(Holding by purpose)
G1587	Anchoring (i.e. holding back)	"
G1597	Supporting = Bearing	
G1599	Shoring = Propping = Strutting	
G1601	Underpinning	
G1603	Suspending = Hanging	
G1613	Storing	
G1623	Maintaining = Preserving *	
G1625	Servicing	
G1635	Stabilising	
G1637	Balancing = Poising	
G1647	Contacting = Touching	
G1649	Adhering = Sticking	
G1651	Cohering	
G1751	Disturbing	
G1753	Agitating = Stirring	
G1755	Striking = Hitting = Tapping = Beating	
G1757	Driving = Ramming	
G1759	Rolling * G2013	
G1761	Sliding	
G1763	Slipping * G2347	
G1765	Rubbing	
G1767	Scrubbing = Scouring	
G1769	Sweeping	
G1771	Grinding * G2651	
G1773	Crushing * G2649	
G1793	Cutting	
G1795	Hewing	
G1797	Sawing	(Cutting by agent)
G1799	Converting = Breaking down	"
G1801	Flat sawing = Plain sawing = Slash sawing = Bastard sawing	"
G1803	Quarter sawing = Quartering = Rift sawing	"

	(Operations)	
	(Disturbing)	
	(Cutting)	
G1813	Thermal cutting (i.e. Heat as agent)	(Cutting by agent)
G1815	Oxygen cutting = Gas flame cutting	(Thermal cutting by heat source)
G1817	Arc cutting	"
G1819	Spark erosion cutting	"
G1821	Electron beam cutting	"
G1831	Incising = Engraving (i.e. Cutting into)	(Cutting by extent of cut)
G1833	Scratching = Scoring	
G1843	Perforating = Piercing	
G1845	Boring	
G1847	Counterboring	(Boring by purpose)
G1849	Reaming	
G1851	Underreaming	
G1853	Countersinking	
G1863	Pressing = Pushing = Thrusting	
G1865	Pulling = Drawing	
G1867	Loading (i.e. Applying load to)	
G1877	Stressing	
G1879	Tensioning = Tensing	(Stressing by force involved)
G1889	Prestressing (i.e. to resist loading)	(Stressing by purpose)
G1891	Pretensioning = Hoyer prestressing	(Prestressing by method)
G1893	Post tensioning = Post stressing	"
G1993	Deforming = Distorting	
G1995	Creeping = Slow deforming	(Deforming by speed)
G1997	Deflecting = Elastic deforming (i.e. Temporary)	(Deforming by permanence)
G1999	Plastic deforming (i.e. Permanent)	"
G2009	Compressing	(Deforming by nature of deformation)
G2011	Compacting	
G2013	Rolling	(Compacting by secondary activity)
G2023	Warping (i.e. Bending/Twisting)	(Deforming by nature of deformation)
G2025	Bending = Flexing = Buckling	
G2027	Folding = Creasing	"
G2029	Wrinkling = Corrugating = Crimping	"
G2039	Straightening	"
G2041	Bowing	(Warping by resultant shape)
G2043	Cupping	X
G2053	Denting = Indenting	(Deforming by nature of deformation)
G2055	Shearing	

(Operations)

G2155	Moving = Flowing	
G2157	Returning	(Moving by nature of motion)
G2159	Circulating = Cirouiting	"
G2161	Rebounding = Bouncing	"
G2163	Reflecting (i.e. rebounding of light)	"
G2173	Pivoting = Swivelling	"
G2175	Rotating = Revolving = Turning *	"
G2177	Spinning	"
G2187	Alternating	"
G2189	Vibrating = Shaking (i.e. alternating on a line)	"
G2191	Rocking	"
G2193	Swinging	"
G2195	Oscillating	"
G2205	Deviating	(Moving by direction)
G2207	Diverging	"
G2209	Radiating	"
G2211	Converging	"
G2221	Ascending = Rising	"
G2223	Descending = Falling	"
G2225	Settling = Subsiding = Sinking	"
G2227	Crossing = Traversing	"
G2237	Penetrating	"
G2239	Permeating	"
G2339	Transferring	
G2341	Placing = Locating = Positioning = Laying	(Transferring by relationship to position)
G2343	Replacing (i.e. transferring to original position)	"
G2345	Displacing = Dislocating	"
G2347	Slipping	"
G2349	Relocating (i.e. transferring to new position)	"
G2359	Orientating	"
G2361	Reorientating	"
G2363	Righting = Erecting	(Orientating by movement involved)
G2365	Inverting	"
G2367	Reversing	"
G2369	Tilting = Tipping	"
G2379	Aligning	(Transferring by relationship to position)
G2381	Centering	

(Operation)

(Transferring)

G2481	Conveying = Carrying = Transporting = Conducting = Bearing *	(Transferring by method)
G2483	Lifting = Raising = Elevating = Hoisting	(Conveying by direction)
G2485	Lowering = Dropping	"
G2495	Impelling = Propelling (i.e. using force)	(Conveying by method)
G2497	Blowing = Blasting	(Impelling by agent)
G2499	Sucking	"
G2509	Sending = Transmitting	(Transferring by material
G2511	Pouring (i.e. Fluids)	transferred)
G2611	(Operations involving general concept of disassembly) Unfixing = Unfastening = Loosening	
G2621	Dividing = Severing	
G2623	Interrupting	
G2633	<del>Breaking</del> = Fracturing	
G2635	Shattering = Fragmenting	
G2637	Bursting = Rupturing	
G2639	Cracking = Splitting = Fissuring = Cleaving	
G2649	Crushing	(Breaking by secondary activity)
G2651	Grinding	"
G2653	Powdering	(Breaking by size of particles)
G2655	Pulverising	"
G2665	Scarifying (i.e. surface)	(Breaking by what broken)
G2667	Ploughing	"
G2677	Blasting = Shotfiring (i.e. explosives)	(Breaking by agent)
G2679	Tearing	
G2779	Separating	
G2781	Disconnecting = Disjoining = Uncoupling = Detaching	(Separating by method)
G2783	Dismantling	
G2793	Sorting	
G2795	Selecting (i.e. by preference)	(Sorting by criteria)
G2797	Classifying (i.e. by similarity)	"
G2799	Grading (i.e. by relative values)	"
G2809	Screening = Sieving = Straining = Sifting	(Separating by method)
G2811	Filtering	"
G2821	Precipitating (i.e. solution)	(Separating by what separated)
G2823	Demulsifying (i.e. emulsion)	

(Operations)

(Operations involving general concept of disassembly)

G2923	Removing = Clearing (100)	
G2925	Depriving = Starving	
G2927	Abrading = Abrasing = Eroding	(Removing by method)
G2929	Scraping	"
G2939	Corroding	"
G2941	Etching	"
G2943	Cleaning = Cleansing	(Removing by what removed)
G2945	Washing = Wet cleaning	"
G2947	Dry cleaning	"
G2957	Purifying = Refining (i.e. impurities)	"
G2959	Sterilising (i.e. micro organisms)	"
G2961	Decolouring = Decolourising	"
G2963	Bleaching	"
G2973	Drying	"
G2975	Dehydrating *G1191	"
G2977	Dewatering	"
G3077	Extracting	
G3079	Emptying = Evacuating = Unloading = Discharging	
G3089	Hollowing = Scooping: cut	(Extracting by method)
G3091	Excavating = Digging	"
G3093	Bottoming up	"
G3095	Dredging (i.e. under water)	(Excavating by physiography)
G3105	Uncovering = Exposing	
G3107	Losing = Giving up = Discharging *	
G3109	Emitting = Giving out	
G3111	Exuding	
G3113	Leaking	
G3115	Expelling = Ejecting	
G3117	Desorbing	
G3127	Distributing = Dispersing = Spreading = Scattering = Diffusing	
G3129	Sprinkling	(Distributing by method)
G3131	Spraying	
G3141	Discarding = Scrapping	
G3143	Wasting	
G3145	Collapsing	
G3147	Telescoping	
G3149	Demolishing = Destroying	

(Operations)

(Operations involving general concept of assembly)

G3249	Fixing = Fastening = Securing	
G3259	Joining	
G3261	Fusing	(Joining by method)
G3263	Welding	"
	(Metal welding)	(Welding by material joined)
G3265	Butt welding	(Welding by form of weld)
G3267	Fillet welding	"
G3269	Plug welding	"
	(Welding without melting)	(Welding by principle used)
G3279	Resistance welding (i.e. electrical resistance)	(Welding without melting by means of plasticising components)
G3281	Resistance butt welding = slow butt welding = Upset butt welding	"
G3283	Flash welding = Flash butt welding	"
G3293	Spot welding = Resistance spot welding	"
G3295	Stitch welding	"
G3297	Multiple spot welding	"
G3299	Series spot welding	"
G3301	Roller spot welding	"
G3311	Seam welding = Resistance seam welding	(Welding without melting by means of plasticising components)
G3313	Butt seam welding = Resistance butt seam welding	"
G3315	Foil butt seam welding = Tape butt seam welding	"
G3325	Projection welding	"
G3327	Percussion welding	"
G3329	High frequency resistance welding = H.F. resistance welding	"
G3339	Forge welding = Fire welding	"
G3341	Blacksmith welding (i.e. manual)	"
G3343	Hammer welding (i.e. mechanical)	"
G3345	Roll welding	"
G3355	Pressure welding	"
G3357	Hot pressure welding	"
G3359	High frequency pressure welding = H.F. pressure welding	"
G3361	Gas flame pressure welding (Synthesise with fuel gas)	"
G3363	Cold welding = Cold pressure welding	"
G3373	Friction welding	(Welding without melting by means of plasticising components)
G3375	Ultrasonic welding	XIV
G3385	Braze welding (i.e. filler metal)	(Welding without melting by agent material)

(Operations)

(Operations involving general concept of assembly)

(Joining)

(Fusing)

(Welding)

(Metal welding)

G3395	Fusion welding (i.e. with melting of components)	(Welding by principle used)
G3397	Arc welding	(Fusion welding by heat source)
G3399	Metal-arc welding (i.e. consumable electrodes)	(Arc welding by electrode type)
G3401	Carbon arc welding	"
G3403	Tungsten arc welding	"
G3405	Inert gas arc welding	(Arc welding by method of shielding)
G3407	Carbon dioxide welding = CO <sub>2</sub> welding	"
G3409	Submerged-arc welding	"
G3411	Electro-gas welding	"
G3413	Vapour shielded welding	"
G3415	Atomic-hydrogen welding	(Arc welding by special processes)
G3425	Gas welding	(Fusion welding by heat source)
-	(Synthesise with fuel gas)	
G3427	Thermit welding	"
G3429	Electron beam welding	"
G3431	Electroslag welding	"
G3441	(Plastic welding i.e. material)	
G3443	Soldering	
G3443	Resistance soldering	(Soldering by heat source)
G3445	Induction soldering	"
G3447	Flame soldering = Torch soldering	"
G3449	Furnace soldering = Oven soldering	"
G3451	Dip soldering	(Soldering by means of introduction of solder)
G3461	Brazing	(Soldering by hardness of solder)
G3463	Saltbath brazing	(Brazing by heat source)
G3465	Flux dip brazing	
G3467	Silver brazing = Silver soldering	(Brazing by brazing material)



(Operations)

(Operations involving general concept of assembly)

(Joining)

G3477	Interlacing	(Joining by method)
G3479	Knotting	"
G3481	Weaving	"
G3483	Binding	"
G3583	Constructing = Assembling = Building	"
G3585	Laminating	(Constructing by what produced)
G3595	Prefabricated constructing	"
G3597	Industrialised constructing = System building	(Constructing by special systems)
G3599	Lift slab constructing	"
G3601	Jack block constructing	"
G3701	Mixing = Blending	"
G3703	Dissolving	"
G3705	Colloiding	"
G3707	Emulsifying	"
G3709	Compounding	"
G3729	Adding	"
G3731	Supplying = Providing	"
G3741	Colouring	"
G3743	Pigmenting	"
G3745	Dyeing = Staining	"
G3747	Moistening = Dampening = Wetting	(Adding by what added)
G3749	Hydrating *G1189	"
G3759	Illuminating = Lighting	"
G3761	Decorating	"
G3763	Veneering	"
G3765	Paperhanging	"
G3767	Graining	"
G3769	Marbling	"
G3789	Inserting (i.e. putting into)	"
G3791	Filling = Loading = Backfilling	"
G3793	Impregnating = Saturating	(Filling by method)
G3795	Soaking = Steeping	"
G3805	Pointing (i.e. joints)	(Filling by what filled)
G3807	Grouting (i.e. joints, cavities)	"
G3809	Dubbing out (i.e. cavities)	"
G3819	Immersing = Submersing = Submerging = Dipping = Sinking *	(Inserting by medium)

(Operations)

(Operations involving general concept of assembly)

G3841	Covering	
G3843	Masking	
G3845	Enclosing (i.e. outer surface)	(Covering by surface involved)
G3847	Encasing = Boxing = Casing	"
G3849	Wrapping = Binding *	"
G3859	Lining (i.e. inside surface)	"
G3861	Paving (i.e. horizontal surface)	"
G3863	Veneering *G3763	(Covering by materials used)
G3865	Paperhanging *G3765	"
G3875	Coating (i.e. cohesive material)	"
G3877	Asphalting	(Coating by material used)
G3879	Painting	"
G3881	Enamelling	"
G3883	Lacquering	"
G3885	Varnishing	"
G3887	Plastering	"
G3889	Pebble dashing	"
G3989	Acquiring = Gaining	
G3991	Sorbing	
G3993	Absorbing	
G3995	Adsorbing	
G3997	Persorbing	
G4017	Concentrating = Collecting = Gathering	
G4019	Coalescing	
G4021	Sintering	
G4023	Flocculating	
G4043	Producing = Generating = Manufacturing = Fabricating = Making	
G4045	Forming = Shaping = Fashioning (i.e. giving form to)	
G4047	Carving (i.e. cutting)	(Forming by secondary activity)
G4049	Turning (i.e. spinning/cutting)	
G4051	Bossing (i.e. striking)	
G4053	Hammering (i.e. striking)	
G6001	Processes	

## APPENDIX B

# NBS

SERVICES

NBS Services  
Mansion House Chambers  
The Close  
Newcastle upon Tyne NE1 3RE

Telephone: 0191 232 9594  
Fax: 0191 232 5714

## FACSIMILE TRANSMISSION

Date: 12th March 1997

To: Douglas Hague

Company:

Fax number: 0115 933 2345

From: John Cann

Message:

Dear Douglas

Thank you for your enquiry concerning the relationship between "Common Arrangement of Work Sections for Building Works (CAWS)" on which our National Building Specification is based and CI/SfB. Also you asked about one of the consultative documents for the new Uniclass classification scheme and about BCIS.

The answer is that there is no official relationship between CAWS and CI/SfB - people tend to use one or the other. The new Uniclass classification scheme includes CAWS as one of its tables, and also includes corresponding tables for each of the tables in CI/SfB, plus some additional tables. However, there will not be, at least in the first edition, a link between the Uniclass CAWS table and, for example, the Uniclass Products table, since it is expected that people will tend to use different Uniclass tables for different purposes rather than using all of them together. As far as BCIS is concerned, again there is no official relationship between this and CAWS. The Uniclass Elements table will cover very similar ground to BCIS.

I hope this has helped to answer your questions - please contact me again if you need any further help.

Regards



John Cann

Number of pages including this one: 1

XIX

Our Ref: PS/TS/MG/jd

17 April 1997

D J Hague Esq MSc MCIOB  
19 Clumber Drive  
Radcliffe-on-Trent  
NOTTINGHAM  
NG12 1DA

12 Great George Street  
Parliament Square  
London SW1P 3AD

Telephone: 0171-222 7000  
Facsimile: 0171-222 9430  
DX: 2348 Victoria 1

Direct line:

Dear Mr Hague

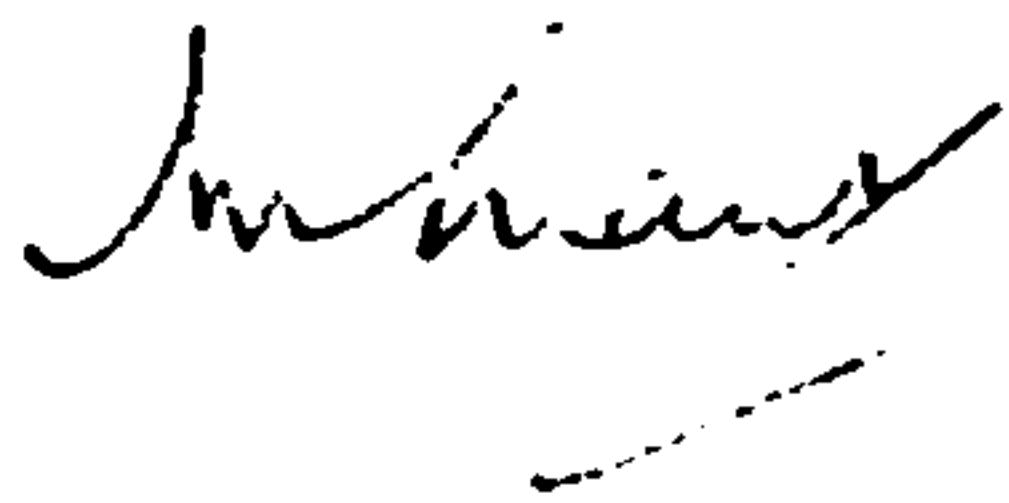
**SMM7**

In reply to your fax letter dated 9 April 1997, addressed to Mr Simon Coates, SMM7 may be copied and used for the purpose for which it was written by any person who wishes to measure or carry out any work or research in connection with it.

It must not, however, be used for profit or reproduced and distributed for gain.

I trust that this answers your query.

Yours sincerely



MICHAEL GRANT  
Technical Services



16 April 1997

D. J. Hague  
19 Colomba Drive  
Ratcliffe-on-Trent  
Nottingham  
NG12 1DA

Southwark Campus  
103 Borough Road  
London SE1 0AA  
Telephone: 0171-815 6601  
Facsimile: 0171-815 6699  
e-mail: john.akeroyd@sbu.ac.uk

Dear Mr Hague

Your letter of 7th April concerning the Construction Industry Thresarus has been passed to me for a reply. I have two comments: -

Firstly the final copy of the CIT was Crown copyright and not that of South Bank University. I am not clear therefore if SBU has any residual rights in this publication and if it is the case that SBU has rights, then I am entirely happy on behalf of the University to agree to you extracting them for you thesis.

In any event my immediate impression would be that quoting even 15 pages of such a publication within a thesis is acceptable practice and does not require any further formalities except that you ensure that you cite the original source.

Yours sincerely

John Akeroyd

cc: P. Noble - Library Manager, Wandsworth Rd

j:2375

XXI

# APPENDIX C

## BUILDING ZONES.

### STRUCTURES

Level 1      Walls                  Partitions                  Foundations

**Explanation.** Elements considered as structural support and division of space

Level 2                  Sub-elements

Classed by **Parts-of** support construction; further defined by **Type-of** material

**Explanation.** Parts function with wall; piers, chimney breasts, arches, copings. Materials, normally, are formed products.

Level 2                  Supported

Cladding                  Finish and Treatment

Classed by **Type-of:** Material or Product

**Explanation.** Cladding and fixings are formed products as purchased. Finishes & Treatments are formless products by definition.

Level 2                  Components

Window                  Door

Classed by **Parts-of** construction and sub-components. Further defined by **Type-of** material and component action.

**Explanation.** Types are classed by action if known (sliding, pivoted, revolving, up-and-over), Parts necessary for function and integrity of component. (Parts of opening classed with 'Sub-elements')

### FLOORS

Level 1                  Floor                  Slab

**Explanation.** Elements considered as platforms for support and horizontal division of internal space. Monolithic and framed construction not differentiated beyond the two designations.

Level 2.                  Sub-elements

Classed by **Part-of** support construction. Further defined by **Type-of** material and product.

**Explanation.** Parts perform a transmission as well as a load-bearing function. Types of materials are processed and formed products as purchased. May also include formless products in slabs, fills below ground slab level and damp proofing.

Level 2                  Supported

Finish and Treatment                  Ceiling

Classed by **Type-of** material or product

**Explanation.** Types of materials and products are as purchased; rigid and flexible sheets, tiles and formless products. Suspended ceilings are included, but not separately classed as a type.

Level 2.                  Stairs

Staircase                  Balustrade

Parts-of complete staircase or barrier, further defined by **Type-of** material.



## Stairs (continued)

**Explanation.** Parts are entities necessary for the integrity and appearance of constructions under the three class headings. Types of style or purpose not classed separately (Open, Dog-leg, Fire escape, Solid, Wreathed).

Level 2.            Components  
                    Stair well        Access

                    Parts-of necessary for second fixings and securing access. Further defined by  
**Type-of material**

## ROOFS

Level 1.            Pitched            Flat

**Explanation.** Elements considered as support for weatherproof surfaces and to enclose space.

Level 2.            Structure  
                    Support            Chimney stack  
                    Parts-of further defined by **Type-of material**

**Explanation.** Parts necessary for integrity of structure and support of weatherproof coverings. Parts of chimney stacks include brickwork, chimney linings and pots

Level 2.            Weatherproofings  
                    Classed by **Type-of material** or product and further defined by **Part-of entity..**

**Explanation.** Materials defined the part as a generic term; tile, slate, asphalt or as an appellation; lead flashing, copper sheeting, mineral felt. Parts include products as purchased and all flashings and gutters except eaves.

Level 2.            Components  
                    Light                Edge                Rainwater disposal  
                    Parts-of further defined by **Type-of material**

**Explanation.** Parts necessary for function and integrity of component and include sub-components as purchased products. Types of lights are not classed, but include dormers, domes, pavement, lantern.

## SERVICES.

### DRAINAGE

Level 1            Above ground                      Below ground

**Explanation:** Systems by the material carried and method. Chambers for access, collection, treatment.

Level 2            Components  
                    Parts-of further classed by **Types-of material** (ceramic, plastic, cast iron)

**Explanation.** Parts; normally, formed products as purchased; fittings to smallest sub-component; formless products (trench fills, surrounds and support).



## SPACE HEATING

Level 1.        Systems  
                  Classed by Type-of fuel.

**Explanation.** Prime fuel at heating source defines System. Electricity used in control and circulation does not form any division.

Level 2.        Components  
                  Producer        Distribution    Delivery  
                  Parts-of necessary for function of system.

**Explanation.** Producers: central and local heating appliances. Distribution: pipework and ducting transporting energy from source. Delivery: radiators, convectors, air distribution outlets. Parts are products and entities as purchased and an integral part of the system.

## EXTERNAL SPACES.

### HARD SURFACES.

Level 1.        Trafficked        Pedestrian

**Explanation.** Spaces adjoining buildings and serving one or more designated purpose. (Access, circulation, assembly, recreation).

Level 2.        Sub-elements  
                  Base                Edge support  
                  Classed by Part-of construction, further defined by Type-of material.

**Explanation.** Parts identified by a structural function; formless products may arise from a process (such as, excavations, demolition). Excludes Surfacing.

Level 2.        Surfacing  
                  Block/slab        Compacted        Applied  
                  Classed by Type-of material and product.

**Explanation.** Products as purchased. Blocks/slab are formed productt of uniform size. Compacted and Applied are formless by definition.

Level 2.        Components  
                  Fixture            Features  
                  Parts-of necessary for function and construction.

**Explanation.** Fixtures include 'Street furniture'. Features include assemblies/constructions with solely an aesthetic function. Parts are sub-components aand products as purchased Excludes SERVICES

## LANDSCAPE

Level 1.        Recreation        Cultivation        Natural feature

**Explanation.** Spaces, within landscape, surrounding or adjoining building. Defined by intended use or dominant physical attribute. (Playing field, garden, woodland)

Level 2.        Shaped ground  
                  Bulk                Surface

Classed by **Type-of material moved**, further defined by a process; excavate, fill, spread. Material may arise from process

Level 2.            Cultivation  
                         Surface            Single and group

Classed by process, further defined by **Type-of plant botanical class** (Tree, shrub, grass)

## DIVIDERS

Level 1.            Wall                    Fence

**Explanation.** Boundaries to define property and separate spaces with different uses.

Level 2.            Sub-element

Classed by **Part-of construction or form**; further defined by **Type-of material**.

**Explanation.** parts identify a function and material types are, normally, formed products.

Level 2.            Components  
                         Access                    Information

**Parts-of necessary fo function**, to smallest purchased product and further defined by **Type-of material**.

**Explanation.** Access may also be used for control purposes. Expressed condition can be met, in certain cases with one sub-component. For example, gosts without a gate = Access without Control.

## APPENDIX D

The forms in this Appendix are reproduced with acknowledgements to Willmott Dixon Maintenance Ltd., the Collaborating Organisation



**WILLMOTT DIXON**

**MAINTENANCE**

**MIDLANDS LIMITED**

**INVOICE**

From: WILLMOTT DIXON MAINTENANCE LIMITED  
1A GROSVENOR ROAD,  
HANDSWORTH, BIRMINGHAM B20 3NP  
TELEPHONE NO: 021 366 1311

INVOICE NUMBER: 657183

DUE FOR PAYMENT: 15/02/92

CUSTOMER ORDER: HPD42161

CUSTOMER: TS0092 CONTRACT: MB5791

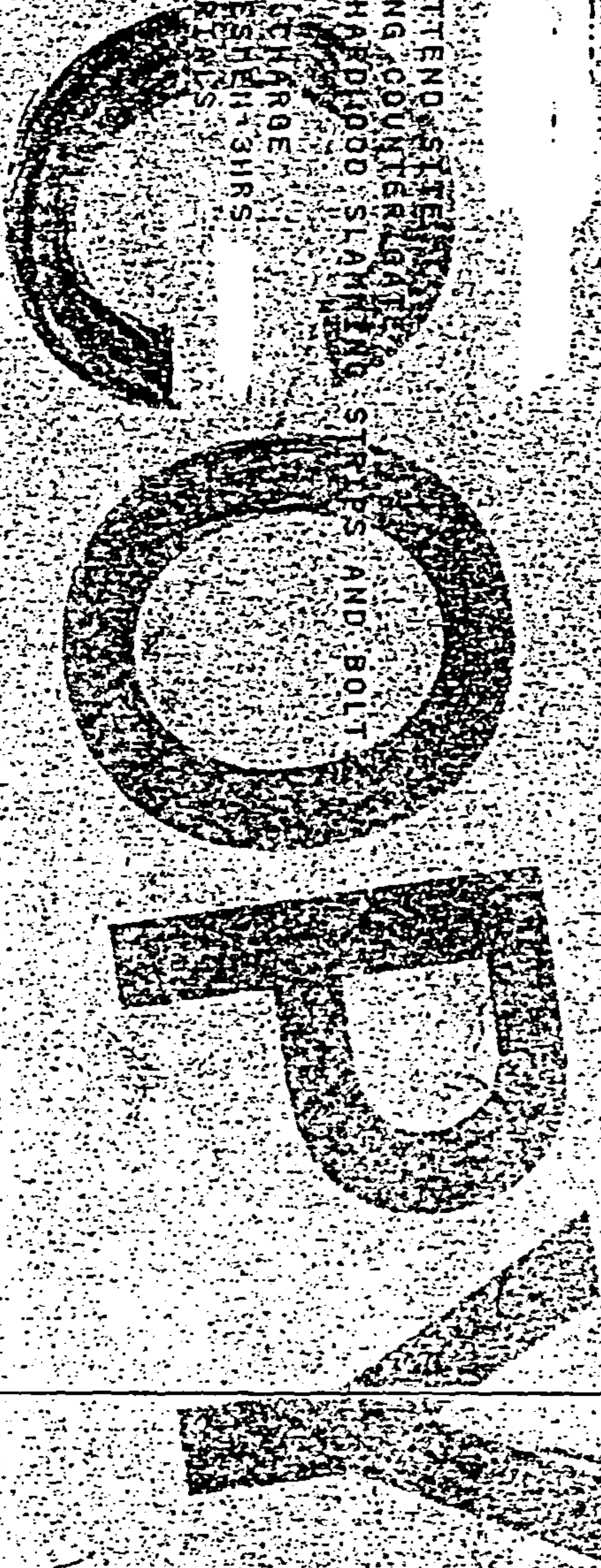
INVOICE/TAX POINT: 16JAH92

PAGE: 1

BATCH NO: 681095

RE:

TO ATTEND SITE  
REHANG COUNTERGATE  
S/F HARDWOOD SLAMING STRIPS AND BOLT  
SITE CHARGE  
TRADERS 3HRS  
MATERIALS



CODE RATE VAT ANALYSIS AMOUNT VAT  
\$ 17.50

GOODS AMOUNT  
VAT AMOUNT

TOTAL DUE

CLIENT <del>XXXXXXXXXX</del>		DATE 12/12/12	<b>DAY WORK SHEET</b>		WILLMOTT DEEDON MAINTENANCE MIDLANDS LIMITED
SITE ADDRESS <del>XXXXXXXXXX</del>		COMPLETE BY 13/12/12	Sanda BEFORE 2PM.		
CUSTOMER CODE	BRANCH No	JOB No	OROCO REF No.	CUSTOMER ORDER No.	
		MB 5701		HPD421611	
WORK REQUIRED	ISSUED TO				

Counter gate needs rehangung & S/F latch.

WORK COMPLETED Rehang gate on helical springs each door  
S/F 3x2 hardwood slamming strip S/F 4" Bolt.

Waiting 3/4 hr for shop to open.

COMMENTS/FURTHER ACTION ETC.	
WORK COMPLETED	
<input checked="" type="radio"/> YES	<input type="radio"/> NO

LABOUR AND TRANSPORT									
DATE	NAME	START TIME	FIN. TIME	HRS	QNO	MATERIALS	MCV	P.CASH	COST
12/12	<del>XXXXXXXXXX</del>	8:00	12:30	2hrs	1	1 @ 3x2 H/WOOD	5		
					10	4" Screws	0		
					1/2	Tube GRIPFILL.	1		
					1/4	BROWN SILICONE.	1		
					1	4" BOLT.	5		

DATE	REG No.	OPEN MILEAGE	CLOSE MILEAGE	HRS TRAVEL	O/No.	PLANT HIRE/SKIPS	OWN.	HIRE	RETURN	COST
12/12	<del>XXXXXXXXXX</del>	32098	32022	2 1/2						

I CONFIRM THAT THE ABOVE WORK HAS BEEN COMPLETED SATISFACTORILY

DATE 13/12/12

SIGNED CLIENT \_\_\_\_\_

SIGNED OPERATIVE \_\_\_\_\_





CLIENT

INPUT COPY



WILLMOTT DIXON  
MAINTENANCE  
MIDLANDS LIMITED

1/2/12

3/12

BEFORE 2PM

CUSTOMER CODE

BRANCH No

JOB No

CROSS REF No

CUSTOMER ORDER No

TEC002

MB 5701

ANDY M

40042616

INV TEXT

① Rehanging counter gate  
S/F hardwood slamingung strips &  
bolt.

657183

	HRS	RATE	SALES VALUE
SITE CHARGE			
TRADESMEN	3		
MATE LABOURER			
ELECTRICAL PLUMBER			
TRADESMEN (PREM)			
MATE/LABOUR (PREM)			
ELECTRICIAN PLUMBER (PREM)			
PLANT			
MATERIALS			
TRANSPORT			
SUB CON			
OTHER			
			INV TOTAL

CLIENT

DATE  
12/12  
COMPLETE BY  
13/12

INVOICE CONTROL  
COPY  
-> BEFORE 2PM !!  
WILLMOTT DODD  
MAINTENANCE  
MIDLANDS LIMITED

CUSTOMER CODE

BRANCH No

JOB No  
MB 5701

CROSS REF No

CUSTOMER ORDER No  
40042161

WORK REQUIRED

ISSUED TO

Counter gate needs rehanging & self latch



**WILMOTT DIXON**  
**MAINTENANCE**  
**MIDLANDS LIMITED**  
**INVOICE**

To: \_\_\_\_\_

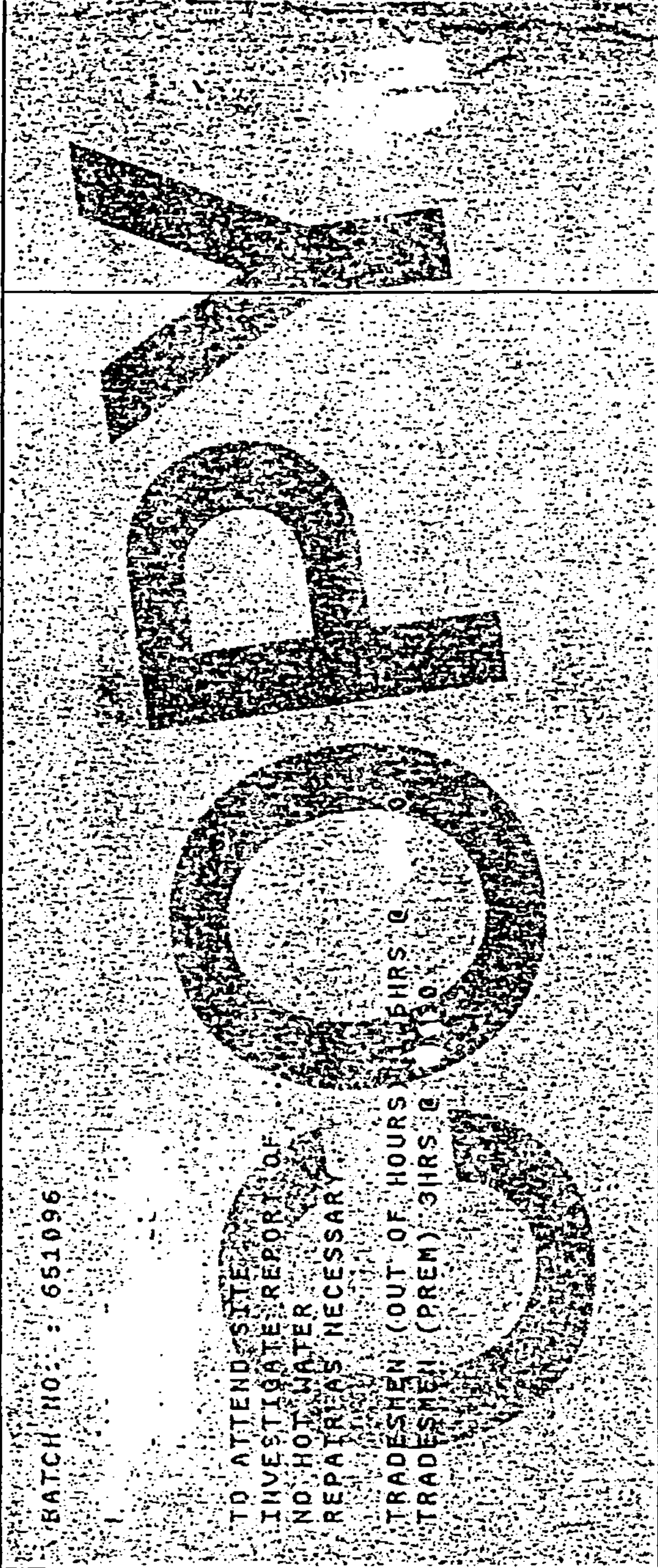
From: WILMOTT DIXON MAINTENANCE LIMITED  
 11A GROSVENOR ROAD,  
 HANDSWORTH, BIRMINGHAM B20 3NP  
 TELEPHONE NO: 021 366 1311

INVOICE NUMBER 657200 DUE FOR PAYMENT 16/02/92 CUSTOMER ORDER  
 CUSTOMER CONTRACT MB5708 INVOICE/TAX POINT 16JAN92 PAGE 1

BATCH NO: 651096

TO ATTEND SITE  
 INVESTIGATE REPORT OF  
 NO HOT WATER  
 REPAIRS AS NECESSARY

TRADESMEN (OUT OF HOURS) 3 HRS @  
 TRADESMEN (PREM) 3 HRS @



CODE RATE AMOUNT VAT  
 S 17 60


GOODS AMOUNT  
 VAT AMOUNT

TOTAL DUE

PLEASE REMIT TO HITCHIN ROAD, SHEFFORD, BEDS, SG17 5JS  
 Registered Office: Hitchin Road, Shefford, Bedfordshire SG17 5JS Telephone: Hitchin (0462) 814455 Fax: (0462) 817255 VAT Registration No. 197 7377 96

CLIENT  
SITE ADDRESS

DATE  
13-12-91  
COMPLETE BY  
13-12-91

**DAY WORK SHEET**  
After 2 p.m.  
  
WILLMOTT DIXON  
MAINTENANCE  
MIDLANDS LIMITED

CUSTOMER CODE	BRANCH No	JOB No	CROSS REF No.	CUSTOMER ORDER No.
	65	MB 5708		

WORK REQUIRED ISSUED TO m.mf OVER PHONE 2:30

No Hot Water. TODAY

WORK COMPLETED  
To check and get hot water working (blockage)

COMMENTS/FURTHER ACTION ETC.  
WORK COMPLETED  
 YES  NO

LABOUR AND TRANSPORT						O/NO	MATERIALS	MCV	P.CASH	COST
DATE	NAME	START TIME	FIN. TIME	HRS						
13/12	m. m	6:15	7:30	1 1/2						

DATE	REG No.	OPEN MILEAGE	CLOSE MILEAGE	HRS TRAVEL	O/No.	PLANT HIRE/SKIPS	OWN.	HIRE	RETURN	COST
13/12		31564	31557	1/4						

I CONFIRM THAT THE ABOVE WORK HAS BEEN COMPLETED SATISFACTORILY  
DATE 13/12/91  
SIGNED CLIENT \_\_\_\_\_  
SIGNED OPERATIVE \_\_\_\_\_




**WILLMOTT DIXON  
MAINTENANCE**

**CONFIRMATION  
NOTE**

CLIENT NAME	OFFICIAL ORDER No. THIS No. MUST BE QUOTED ON THE INVOICE
ADDRESS	
DATE 12.12.9.	

DESCRIPTION OF JOB NO HOT WATER.  
 MAKE SURE STOP TAP ON COLD WATER SUPPLY OPEN  
 MAKE SURE COLD HEATED TANK FILLING UP REGULATE  
 BALL VALVE TO FILL. CONCLUDE AIR LOCK TAKE OFF  
 INWARDS TO 2ND TAPS IN BATHROOM AND 1ND TAP IN  
 KITCHEN. ALSO TAKE OFF H. WATER PIPE TO WASHING  
 MACHINE TO BLEED. NO GOOD TRY TO BLOW BACK TO  
 HEATED. FIND OUT BACK BOILER RECENTLY INSTALLED  
 BUT NO VISIBLE DRAIN OFF OR BLEED TAP AIR LOCK.

IS INSIDE. CANNOT TAKE OUT GAS FIRE. HAVE NOT BEEN  
 ON ACCORDS COURSE

TRADE	NAME	TIME ON	TIME OFF	TOTAL HRS.
PLUMBER.	A.D.	9:00 PM	12:00 PM	3

GENERAL COMMENT TOTAL

SOME-BODY MUST GO TOMMOROW AFTER 2 PM

CLIENT AUTHORISATION	COMPANY STAMP	TO BE RETAINED
SIG. ....		
NAME. ....		
JOB TITLE. ....		





CLIENT

# INPUT COPY



WILLMOTT DIXON  
MAINTENANCE  
MIDLANDS LIMITED

13-12-91

CUSTOMER CODE

BRANCH No.

JOB No

CROSS REF No.

CUSTOMER ORDER No.

10001

65

MB 5708

## INV TEXT

①

②

No hot water.  
Repair as necessary

657200

	HRS	RATE	SALES VALUE
SITE CHARGE			
TRADESMEN			
MATE LABOURER			
ELECTRICAL PLUMBER			
TRADESMEN (PREM)			
MATE/LABOUR (PREM)			
ELECTRICIAN PLUMBER (PREM)	3		
PLANT			
MATERIALS			
TRANSPORT			
SUB CON			
OTHER			
			INV TOTAL

(1 1/2 @ -

(3 @ -

out of hrs. 1 1/2



## APPENDIX E

The extracts of Prolog program listings were written by Dr. P. R. Innocent and reproduced with his permission.

```

REM A program module to test for unwanted punctuation after a word and
REM then replace with a space.
REM Adapted from James, M (1992) 'QBasic The language of MS DOS' Leyburn,
REM I/O Press, pp.83, 84
CLS
PRINT : PRINT : LINE INPUT "Convert to Prolog--", sS
cS = ""
FOR i% = 1 TO LEN(sS)
  cS = MIDS(sS, i%, 1)
  IF INSTR(1, ",;:!?.", cS) <> 0 THEN
    MIDS(sS, i%, 1) = CHR$(0)
  END IF
NEXT i%
PRINT : PRINT : PRINT "[" + LCASES(sS) + "]"
REM ascii for a space is 32; 58 = colon; 59 = semi-colon; 0-9 = 30-39
' uppercase A-Z = 65 - 90 and lowercase a-z = 97 - 122; other char
' < = > ? @ are 60 - 64 incl.

```

```

DECLARE SUB result (sS)
DECLARE SUB blanko (sS)
REM another version of PROLCHOP.PRO intended to separate words in a string
'with comma and trailing space.

```

```

CLS
PRINT : PRINT : LINE INPUT "Type description--", sS
cS = ""
FOR i% = 1 TO LEN(sS)
    cS = MIDS(sS, i%, 1)

    IF ((ASC(cS) <= 48) OR (ASC(cS) >= 58) AND (ASC(cS) <= 64) OR (ASC(cS) >=
91) AND (ASC(cS) <= 96) OR (ASC(cS) >= 122)) THEN
        MIDS(sS, i%, 1) = " "
    END IF

NEXT i%
CALL blanko(sS)

```

```

SUB blanko (sS)
'DO

```

```

l% = LEN(sS)
REM Test for trailing blank
REM @@@@ @@@@ @@@@ @@@@ @@@@
IF MIDS(sS, l%, 1) = " " THEN
sS = (LEFTS(sS, (l% - 1)))
END IF
aS = " " + " "
a% = INSTR(1, sS, aS)
l% = LEN(sS)
w% = (a% - 1)
e% = l% - a%
'IF a% = 0 THEN
'CALL result(sS)
'END IF

sS = (LEFTS(sS, w%) + RIGHTS(sS, e%))
'LOOP UNTIL a% = 0

```

```

CALL result(sS)
END SUB

```

```

SUB result (sS)
PRINT : PRINT "result = "; LCASES(sS)

```

```

END SUB

```

```

'A test program to read data and print it
'Called DATEST2.PRO (DATATEST.PRO remains on A:\RESEARCH)
CLS
PRINT : PRINT
LINE INPUT "Enter sentence at ? "; AS
DJ = 14
DIM wordS(DJ): DIM JU(DJ)
FOR S = 1 TO DJ
  READ wordS(S)
  READ JU(S)
  wordS = wordS(S)
  JU = JU(S)
  BS = LCASES(AS)
  P = INSTR(BS, wordS)
  IF P >= 1 THEN
    PRINT "- "; " I found "; wordS; " at "; P; " on loop "; S; " "; JU
    L = L + S / S
    T = (JU + T)
    restS = LEFTS(BS, P - 1 + LEN(wordS))
    PRINT restS
    PRINT
  END IF
NEXT S
RESTORE

PRINT : PRINT "Total JU = "; T; " and a keyword occurred "; L; " times."
AV = T / L: MP = AV / 32
PRINT : PRINT "Average value = "; AV; " and the probability value is "; MP
IF MP > .99 THEN
  fuzzS = "certain"
END IF
IF MP <= .99 AND MP >= .62 THEN
  fuzzS = "very likely"
END IF
IF MP <= .619 AND MP > .45 THEN
  fuzzS = "more likely to be about drains than"
END IF
IF MP <= .45 THEN
  fuzzS = "quite unlikely"
END IF
PRINT : PRINT "The input line is "; fuzzS; " to refer to roof work."

DATA roof,32, door,8, ladder,30
DATA lock,6, drain,19, "slipped slates",31
DATA jamb,4, rod,18, manhole,17, clear,15
DATA blockage,20, ridge,34, eaves,35, fascia,36
REM This program prints the input line (or rather, leaves it on screen),
REM prints, for each keyword matched by P = INSTR(AS,wordS),
REM '- I found wordS at P at loop Nr. S'
REM and data sequence Nr. After the loop ends, the program stops with the
REM words "Total JU = X and a keyword occurred L times.'
REM 11/5/95- Lines added to average integer data values and divide result

```



```

CALL trimstring(BS, NS)
fxS = dS + "\" + NS + ".TXT"
CLS
PRINT fxS; " is the file being 'presented' for your inspection."
PRINT "Note that keywords, when identified, are shown at position P and"
PRINT "printed before the line of text in which they have been found."
PRINT "Although this is not invariably so after amendments on 24/5/96"

CALL dividingLine
CALL getafile(fxS)

PRINT : PRINT " Total JU "; AV; " Total L "; L
LOOP UNTIL (EOF(2))
CLOSE #2
DATA wall,10,door,2>window,3>felt,4>chimney,5>flue,6>fireplace,7>jamb,8
DATA transom,9>transome,10>roof,51>eaves,32>gutter,33>ladder,34>steps,15
DATA ceiling,16>closer,17>glass,18>Yale,19>cylinder,20>latch,21>beads,22
DATA floor,23>joists,24>undercoat,25>sweep,26>rubbish,27>leaves,28
DATA manhole,19>rod,30>unblock,17>"layers of felt",32>"pan",33>concrete,34
DATA track,35>leak,36>handle,37>"panel",38>"wc pan",39>wc,39>flush,41
DATA flashing, 52, "flash", 53, pointing,34
END
REM This version amended slightly 5/8/97 in search of solution found before
'writing thesis, but since lost in development. The calculation of average
'for the probability function needs adding (working algorithm in D2TATEST.PRO

SUB dividingLine
PRINT
PRINT
"*****"
PRINT
END SUB

SUB eight (BS, NS)
NS = LEFTS(BS, 8)
END SUB

SUB five (BS, NS)
NS = LEFTS(BS, 5)
END SUB

SUB four (BS, NS)
NS = LEFTS(BS, 4)
END SUB

SUB getafile (fxS)
OPEN fxS FOR INPUT AS 5
OPEN "SCRN:" FOR OUTPUT AS 6
DO
LINE INPUT #5, CS

```

```

PRINT #6, LCASES(CS)
  CALL loadKeyword(wordS, JU, CS, P, S, L, T, AV, MP)
  LOOP UNTIL (EOF(5))
CLOSE #5: CLOSE #6
PRINT "Average of JU"; AV; " and ";
SLEEP
REM This subroutine opens a job file for scanning, using another
REM subroutine, loadKeyword, to match keywords with INSTRS function.
END SUB

```

```

SUB loadKeyword (wordS, JU, CS, L, T, AV, P, S, MP)

```

```

REM This sub is adapted from DOCTOR in Krutch, J.
'Experiments in Artificial Intelligence, pp. 125-6
'24th May 1995 was adapted into ACCESS20.PRO using
'program called D2TATEST.PRO

```

```

DJ = 44: REM **Remember to check DATA words****
DIM wordS(DJ): DIM JU(DJ)
FOR S = 1 TO DJ
  READ wordS(S)
  READ JU(S)
  wordS = wordS(S)
  JU = JU(S)
  BS = LCASES(CS)
  P = INSTR(BS, wordS)
  IF P >= 1 THEN
    T = T + JU
    PRINT ":- "; "Word "; wordS; " at "; P; " JU= "; JU; "Total "; T
    L = L + S / S
    AV = T / L
    PRINT "Line value of JU = "; AV
  END IF

```

```

'A select on CASE to send text file to relevant directory for Prolog
' SELECT CASE AV

```

```

' CASE IS < 30
' CALL strFlr

```

```

' CASE 30 TO 70
' CALL strRoof
' END SELECT

```

```

NEXT S
RESTORE

```

```

REM The foregoing SELECT asks for one of near identical subroutines and
'the transfer of a meaningful AV when called.

```

```

END SUB

```

```
SUB seven (BS, NS)
NS = LEFTS(BS, 7)
END SUB
```

```
SUB six (BS, NS)
NS = LEFTS(BS, 6)
END SUB
```

```
SUB strFlr
CHDIR "C:\HAGUE\FLOORS"
SHELL "COPY fxS C:\HAGUE\FLOORS"
END SUB
```

```
SUB strRoof
CHDIR "C:\HAGUE\ROOFS"
SHELL "COPY fxS C:\HAGUE\ROOFS"
```

```
END SUB
```

```
SUB trimstring (BS, NS)
'subroutine to check for position of spaces in the file names listed in
'a DOS directory (maximum 8 characters long).
'Willmott-Dixon never identified a job with a code less than 4 alphanumeric
```

```
zS = MIDS(BS, 5, 1)
IF ASC(zS) = 32 THEN
T% = 1
ELSE T% = 0
END IF
```

```
REM filename is 4 char long
```

```
ziS = MIDS(BS, 6, 1)
IF ASC(ziS) = 32 THEN
ti% = 1
ELSE ti% = 0
END IF
```

```
REM filename is 5 char long
```

```
zjS = MIDS(BS, 7, 1)
IF ASC(zjS) = 32 THEN
tj% = 1
ELSE tj% = 0
END IF
```

```
REM filename is 6 char long
```

```
zkS = MIDS(BS, 8, 1)
IF ASC(zkS) = 32 THEN
tk% = 1
ELSE tk% = 0
END IF
```



REM filename is 7 char long

```
zls = MIDS(BS, 9, 1)
IF ASC(zls) = 32 THEN
  tl% = 1
ELSE tl% = 0
END IF
```

REM last space before TXT file extension as listed in the directory

'This next algorithm from JAMES, M. (1992) "QBASIC" Leyburn, I/O Press,p.164  
'calls up another subroutine that properly identifies the file for DOS to open

C = (T% + ti% + tj% + tk% + tl%) 'Value a maximum of 5 spaces

```
IF C = 1 THEN
  CALL eight(BS, NS)
```

```
ELSEIF C = 2 THEN
  CALL seven(BS, NS)
```

```
ELSEIF C = 3 THEN
  CALL six(BS, NS)
```

```
ELSEIF C = 4 THEN
  CALL five(BS, NS)
```

```
ELSEIF C = 5 THEN
  CALL four(BS, NS)
```

```
ELSEIF C = 0 THEN
```

```
  NS = BS
```

```
END IF
```

END SUB

```

'A test program to read data and print it
'Called D2TATEST.PRO (DATATEST.PRO remains on A:\RESEARCH)
CLS
PRINT : PRINT
INPUT AS
DJ = 23
DIM word$(DJ): DIM JU(DJ)
FOR S = 1 TO DJ
  READ word$(S)
  READ JU(S)
  word$ = word$(S)
  JU = JU(S)
  BS = LCASE$(AS)
  P = INSTR(BS, word$)
  IF P >= 1 THEN
    PRINT "- "; " I found "; word$; " at "; P; " on loop "; S; " "; JU
    L = L + S / S
    T = (JU + T)
  END IF
NEXT S
RESTORE

PRINT : PRINT "Total JU = "; T; " and a keyword occurred "; L; " times."
AV = T / L: MP = AV / 32
PRINT : PRINT "Average value= "; AV; " and the probability is "; MP;
PRINT "where unity is certainty."

IF MP >= .62 THEN
  fuzz$ = "very likely"
END IF
  IF MP <= .619 AND MP > .45 THEN
    fuzz$ = "more likely to be about drains than"
  END IF
  IF MP <= .45 THEN
    fuzz$ = "quite unlikely"
  END IF
  PRINT : PRINT "The input line is "; fuzz$; " to refer to roof work."

DATA roof,32, door,8, ladder,30
DATA lock,6, drain,19, "slipped slates",31
DATA jamb,4, rod,18, manhole,17, clear,15
DATA clip, 18, "molten pitch", 30, dormer, 32, ridge, 28
DATA hip, 30, "split felt", 28, rosemary, 28, verge, 30, eaves, 32
DATA rafter, 32, gutter, 30, "down pipe", 30, "nail tile", 32
REM This program prints the input line (or rather, leaves it on screen),
REM prints, for each keyword matched by P = INSTR(AS,word$),
REM '- I found word$ at P at loop Nr. S'
REM and data sequence Nr. After the loop ends, the program stops with the
REM words 'Total JU = X and a keyword occurred L times.'
REM 11/5/95- Lines added to average integer data values and divide result
REM by 32 for the scale of probability and print the results. This was further
REM developed with another IF THEN to measure the likelihood of the input
REM referring to roof work.

```

## Appendix

This Appendix contains details of the prolog program and data files used to parse preprocessed descriptions of work carried out.

5 The result of preprocessing the raw data is the generation of a file called 'sentences' which contains separate lines each in prolog fact format as follows:

sentence(N,P).

where N is the sequential number of the sentence (line) in the file and P is a list of sentence words as follows:

P=[this,is, a,sentence]

The output of the prolog program is written to the console (terminal screen) and consists of the result of parsing each sentence grammatically.

The following is a sample of the sentences file which contains:

```
sentence(1,['To',attend,site,and,rod,all,manholes,on,drain,run,to,clear,blocked,'WC'
]).
sentence(2,['To',attend,site,and,rod,all,holes,on,drain,run,to,clear,blocked,'WC']).
```

The program file is 'hague1' which contains all the top level predicates to parse the sentences.

The following output is generated by the program when it is run:

```
/* run of test program and code follows */
BNR prolog Version 3
```

```
?- test.
```

```
sentence no:1
```

```
verb_phrase(verb(rod),noun_phrase(determiner(all),noun(manholes)))
```

```
aux_phrase1(preposition(on),qual_noun(qualifier(drain),noun(run)))
```

```
goal_phrase(infinite_verb(infinite(to),verb(clear)),aux_phrase2(preposition(blocked),n
oun(WC)))
```

```
[]
```

```
sentence no:2
```

```
verb_phrase(verb(rod),noun_phrase(determiner(all),noun(holes)))
```

```
aux_phrase1(preposition(on),qual_noun(qualifier(drain),noun(run)))
```

```
goal_phrase(infinite_verb(infinite(to),verb(clear)),aux_phrase2(preposition(blocked),n
oun(WC)))
```

```
[]
```

```
?- test().
```

```
YES
```

```
?-
```

```
*/
```

The program has been written as a demonstrator to show how a general approach can be taken to parsing phrase structured grammars in a limited context. Although not used, semantic checks have been included for demonstrating that they can be included. The semantic checks for each grammar context are

written to always succeed but could easily be modified to check that certain conditions are satisfied so that the parsed words roles are compatible in a real world interpretation.

Generality has been addressed by allowing the language rules to be accessed via a variable whose value could be set according to a simple word filter from a string in the sentence (preprocessing stage).

Similarly the 'classes' file could be dynamically selected according to the outcome of a simple filter.

The program is not commented as a working knowledge of prolog is necessary to understand fully the semantics of the program.

To run, start BNR prolog and consult 'hague1' then enter the query 'test'. If this fails check that you have the files 'sentences', 'class' (containing a list of synonyms for a given class - currently only for drains), 'Edinburgh' (used to make BNR prolog look like Edinburgh syntax), 'drains', (containing the lexicon for drains work), and 'grules\_h' (containing the grammar rules for parsing phrases).

These files can be generated by editing the following text.

Please note that to run in other prolog systems modification are needed - some serious!

```
/* main file - top level Hague1 */
/* initialise - loads necessary grules interpreter */
```

initialise:-

```
[retractall(sentence(,_));true],
consult('Edinburgh'),
consult('sentences'),
consult('classes').
```

```
/* test predicate just for basic debugging - it would need
development to allow file input in pdc prolog and note that the predicate 'once' is not in
pdc prolog*/
```

test:-

```
once(initialise),
Rules = 'grules_h',
sentence(Num,String),
nl,write('sentence no:',Num),
once(dosentence(Rules,String)),fail.
```

test.

dosentence(Rules,String):-

```
classify(String,Classes),
selecta(Class,Classes),
dcg_load(Rules),
dcg_load(Class),
```

```
%get semantics and lexicon (not grammar rules)
stripline(String,P),
analyse(P,S,Rest),
prettyprint(S),
nl, write(Rest).
```

```
/* simple main control loop that builds up the list of phrases in the input */
analyse([],[],[]).
```

```

analyse(P,[Structure|T],Rest):-
    phraseanal(Structure,P,L),
    analyse(L,T,Rest).
analyse(L,[],L).

/* classify string into classes */
classify(String,Classes):-
    classes(Classes),
    clasif(String,Classes,Classes).
clasif(_,[],[]).
clasif(String,[Class|Classes],[Class|Others]):-
    member(Class,String),
    clasif(String,Classes,Others).
clasif(String,Classes,[Class|Others]):-
    clasif(String,Classes,Others).

/* select a class */
selecta(Class,[Class|_]). %just take the first in this demo

/* stripping the terms from the input - replace by pdc prolog stuff */
stripline(P,P1):-
    append(['To'],X,P),
    append(Z,['and'|P1],X),
    ([not(member('Mr',Z))];[not(member('Mr.',Z))]).

stripline(P,P1):-
    append(['To'],X,P),
    (append(Q,['.',],Q1);append(Q,['.',],Q1)),
    append(Q1,P1,X),
    (member('Mr',Q1);member('Mr.',Q1)).
stripline(X,X).

/* print routine */
prettyprint([]).
pretty_print([H]):- nl,write(H).
prettyprint([H|T]):-
    nl,mywrite(H),
    prettyprint(T).
mywrite(H):-
    nl,write(H).

/* utility predicates */

member(X,[X|_]).
member(X,[_|P]):-member(X,P).

append([],X,X).
append([X|T1],T2,[X|T3]):- append(T1,T2,T3).

/* end of program hague1 */

/* File 'drain'
a lexicon for drains */
/* lexicon for drains */
infinite(infinite(to),[to|Rest],Rest).
qualifier(qualifier(drain), [drain|Rest],Rest).
determiner(determiner(all),[all|Rest],Rest).
noun(noun(manholes), [manholes|Rest],Rest).

```

```

noun(noun(holes), [holes|Rest],Rest).
noun(noun(run),[run|Rest],Rest).
noun(noun('WC'), ['WC'|Rest],Rest).
verb(verb(rod),[rod|Rest],Rest).
verb(verb(clear),[clear|Rest],Rest).
preposition(preposition(on),[on|Rest],Rest).
preposition(preposition(blocked), [blocked|Rest],Rest).

```

```

/* semantic check for drains */
semantic_check(,_).

```

```

/* end of file 'drain' */

```

```

/* grammar rules follow */

```

```

/* File 'grules_h' */

```

```

phraseanal(VP) --> verb_phrase(VP).
phraseanal(VP) --> aux_phrase(VP).
phraseanal(VP) --> goal_phrase(VP).
verb_phrase(verb_phrase(Verb,NP)) -->[ verb(Verb),noun_phrase(NP),
    {semantic_check(Verb,NP)} ].
noun_phrase(noun_phrase(Det,Noun)) --> [determiner(Det), noun(Noun),
    {semantic_check(Det,Noun)}].
aux_phrase(aux_phrase1(Prep,NP)) -->[ preposition(Prep),qualified_noun(NP),
    {semantic_check(Prep,NP)}].
aux_phrase(aux_phrase2(Prep,NP)) --> [preposition(Prep),noun(NP),
    {semantic_check(Prep,NP)}].
goal_phrase(goal_phrase(Inf,Aux)) --> [infinite_verb(Inf),aux_phrase(Aux),
    {semantic_check(Inf,Aux)}].
qualified_noun(qual_noun(Q,N)) --> [qualifier(Q),noun(N),
    {semantic_check(Q,N)}].
infinite_verb(infinite_verb(Inf,Verb)) --> [infinite(Inf),verb(Verb)].
/* end of file grules_h */

```

```

/* file of class information */
classes([drain,drains,roofs]).
/* end of classes file */

```

## References

- ALLAN, F. G. (1966) Comparative estimating. *Work study & management services*, (pt. Oct.), pp. 533-4
- ASHFORD, J. (1977) *Statistics for management*. London, Institute of Personnel Management.
- BS 3138:1992. *Glossary of terms used in management services*. London, BSI.
- BUILDING RESEARCH STATION (1959) *Principles of modern building*. 3rd ed. 1, London, HMSO.
- BULFINCH, T. (1963) *Mythology*. London, Spring Books.
- CAMPBELL, J. (1984) *Grammatical man*. Harmondsworth, Pelican Books.
- CAREY, S. (1978) The child as word learner. In: HALLE, M. et al. eds., *Linguistic theory and psychological reality*, Cambridge, Mass., MIT Press.
- CHOMSKY, N. (1988) *Language and problems of knowledge: the Managua lectures*. Cambridge, Mass., MIT Press.
- CIB COMMISSION W31 (1972) *CIB Master lists*. Report No.18, Rotterdam, International Council for Building Research Studies and Documentation-CIB.
- CLOCKSIN, W. F. and MELLISH, C. S. (1987) *Programming in Prolog*. 3rd ed., Berlin, Springer-Verlag.
- COORDINATING COMMITTEE for PROJECT INFORMATION (1987) *Common arrangement of work sections for building work*. Newcastle-upon-Tyne; Building Project Information Committee.
- ELLIOTT, D. A. (1982) Schedules of rates for local authority building maintenance work. HARLOW, P. ed. *Technical Information Service*, No. 8, Ascot, Chartered Institute of Building.
- FABB, N. (1994) *Sentence structure*. London and New York, Routledge.
- FIEBLEMAN, J. K. (1954) Theory of integrative levels. *British journal for the philosophy of science*. 5, pp. 59-66
- FODOR, J. A. (1980) Fixation and belief and concept acquisition. In: PIATTELLI-PALMARINI, M. ed. *Language and learning: the debate between Jean Piaget and Naom Chomsky*, Cambridge, Mass., Harvard University Press.
- FODOR, J. A. et al. (1980) Against definition. *Cognition*, 8, pp. 263-367
- FORSYTH, R. (1978) *The BASIC idea*. London, Chapman and Hall.

- GOLDS, R. N. (1990) Pitfalls and auditing of measured term contracts. In: *Measured term contracts, Building Maintenance Information Special report*, Serial 195, London, Building Maintenance Information Ltd., pp. 32-35
- HAGUE, D. J. (1977) *Incentives in building maintenance using comparative estimating* MSc., University of Manchester Institute of Science and Technology.
- \_\_\_\_\_. (1983) *An investigation into the type, durations and frequency of work tasks in building maintenance* An unpublished report for the Chartered Institute of Building, Queen Elizabeth II Silver Jubilee Scholarship and Award Scheme.
- HOBBS, S. J. (1978) Collection and use of building maintenance cost data. *Current papers*, CP 3/78, (Oct.), Watford, Building Research Establishment.
- HOLMES, R., DROOP, C. and MELLOR, P. (1985) A coding system for building maintenance. HARLOW, P. ed. *Technical Information Service*, No. 47, Ascot, Chartered Institute of Building.
- \_\_\_\_\_. and MELLOR, P. (1985) Maintenance coding and monitoring: two case studies. HARLOW, P. ed. *Technical Information Service*, No. 53, Ascot, Chartered Institute of Building.
- \_\_\_\_\_. (1987) Management of feedback information. In: ISRAEL INSTITUTE OF TECHNOLOGY: FACULTY OF CIVIL ENGINEERING, 2, *Proceedings of the 4th International Symposium on Robotics and Artificial Intelligence in Building Construction*, (June), Haifa, Israel Institute of Technology, pp. 639-656
- HURLSTON, C. I. (1967) Second generation analytical estimating. *Work study & management services*, (pt. Oct.), pp. 566-8.
- JAMES, M. (1984) *Artificial intelligence in BASIC*. Sevenoaks, Newnes Technical Books.
- \_\_\_\_\_. (1985) *Classification algorithms*. London, Collins.
- JOHNSON-LAIRD, P. N. and QUINN, J. G. (1976) To define true meaning. *Nature*, 264, pp. 635-6
- \_\_\_\_\_. (1983) *Mental models*. Cambridge, Cambridge University Press.
- KELLY, T. (1826) The practical builder's perpetual price-book. In: KELLY, T. *Practical carpentry, joinery and cabinet making*, London, J. Rider.
- KRUTCH, J. (1986) *Experiments in AI for microcomputers*. 2nd ed., Indianapolis, Sams.
- LEECH, G. (1974) *Semantics*. Harmondsworth, Penguin Books.
- McKINNON, B. (1971) Construction maintenance repair; the universal yardstick. *Business administration*, (pt. April), pp. 44-7



MILLER, G. A. ed. (1967) *Psychology of communication*. 1975 ed. Basic Books, pp. 125-167

NAGAO, M. (1989) *Machine translation*. Oxford, Oxford University Press.

NATIONAL BUILDING SPECIFICATION. Unified Classification for the Construction Industry (1995) *Consultative document 8. Building elements. Draft for comment*. (Sep.), Newcastle-upon-Tyne, NBS Services.

NSR MANAGEMENT, (1995) *National Schedule of Rates*. Aylesbury, NSR Management.

PERRY, G. (1993) *QBasic by example*. Indianapolis, Que Corporation.

PILCHER, R. (1992) *Principles of Construction Management*, 3rd ed., London, McGraw Hill.

PUTNAM, H. (1975) The meaning of "meaning". In: GUNDERSEN, K. ed *Language, mind and knowledge*, 7, Minnesota, University of Minnesota Press, pp.131-193.

RANDALL, R. H. (1978) Building defects and feedback in PSA. HARLOW, P. ed., *Maintenance Information Service*, No. 4, (Summer), Ascot, Institute of Building.

RICHMOND, P. (1965) Transformation and organization of information content: aspects of recent research in the art and science of classification. In: *The Library Association. Research pamphlet No. 1: Classification and information control*, London, The Library Association.

ROBERTS, M. et al. (1970) *Thesaurus for the construction industry*, for the North Western Polytechnic School of Librarianship and the Brixton School of Building, First draft, (November), unpublished work. (See Appendix B, letter from P. Noble, South Bank University).

ROUGIER, P. and LEFFLEY, M. (1994) The diagnosis of building defects by computer. HARLOW, P. ed. *Construction papers*, No. 34, Ascot, Chartered Institute of Building.

ROYAL INSTITUTE OF BRITISH ARCHITECTS (1961) *SfB/UDC Building filing manual*. London, RIBA Technical Information Service.

\_\_\_\_\_. (1968) *Construction indexing manual*. London, RIBA Publications Ltd.

ROYAL INSTITUTION OF CHARTERED SURVEYORS and BUILDING EMPLOYERS CONFEDERATION (1988) *Standard method of measurement of building works*, 7th. ed. London, Royal Institution of Chartered Surveyors and Building Employers Confederation.

RUBERG, K. et al. (1989) House calls: building and maintaining a rule-base. *Knowledge acquisition*, 1, pp. 379-401

- RUSSELL, B. (1948) *Human knowledge: its scope and limits*. London, Geo. Allen & Unwin Ltd.
- SAVILLE, A. (1990) Computerised maintenance management in a housing authority *Technical Information Service*, No.117, HARLOW, P. ed. Ascot, Chartered Institute of Building.
- SCHLESINGER, I. M. (1995) On the semantics of the object. In: AARTS, B. and MEYER, C. F. eds. *The verb in contemporary English*, Cambridge, Cambridge University Press, 4, pp. 54-74
- SHARP, J. R. (1965) *Some fundamentals of information retrieval*. London, Andre Deutsch.
- Shorter OED. (1944) 3rd. ed., revised and corrected, 1978.
- SKINNER, N. P. and KROLL, M. E. (1984) Maintenance feedback. In: HARLOW, P. ed. *Managing building maintenance*, Ascot, The Chartered Institute of Building, pp. 53-62
- TOMLINSON, H. (1969) Report on work for new general classification scheme. In: *The Library Association Research Pamphlet No. 1: Classification and Information Control*, London, The Library Association, pp. 29-41.
- VICKERY, B. C. (1960) *Faceted classification*. London, Association of Special Libraries and Information Bureaux.
- WINOGRAD, T. (1973) A procedural model of language understanding. In: *Computer models of thought and language*. SCHANK, C. R. and COLBY, K. M. eds., San Fransisco, W. H. Freeman & Co.
- WYKES, T. and JOHNSON-LAIRD, P. N. (1977) How do children learn the meaning of verbs? *Nature*, 268, pp. 326-7

## Bibliography

- AOUD, G. et. al. (1995) The conceptual modelling of construction management information. In: *Automation in construction*, 3, pp. 267-282
- ANDERSON, J. R. (1984) Spreading activation. In: ANDERSON, J. R. and KOSSLYN, S. M. eds. *Tutorials in learning and memory*, San Fransisco, Freeman. pp. 61-90
- ASHWORTH, A. (1986) Price books and schedules of rates. HARLOW, P. ed. *Technical Information Service*, No. 65, Ascot, Chartered Institute of Building.
- BARBER, C. (1993) *The English language: a historical introduction*. Cambridge, Cambridge University Press.
- BARR, A. and FIEGENBAUM, E. A. eds. (1985) Semantic primitives. In: *The handbook of artificial intelligence*, 1, ch. III, sect. C6, Reading, Mass., Addison-Wesley Publishing Company Inc. pp. 207-215
- BAXENDALE, A. T. (1987) A coding system for new building work. HARLOW, P. ed. *Technical Information Service*, No.86, Ascot, Chartered Institute of Building
- BRUCE, D. J. (1956) Effects of context upon intelligibility of heard speech. In: CHERRY, C. ed. *Information theory*, London, Butterworths, pp. 245-252.
- BURCHFIELD, R. W. ed. (1996) *The new Fowler's modern English usage*. 3rd ed. Oxford, Clarendon Press.
- BUTLER, C. S. ed. (1992) *Computers and written texts*. Oxford, Basil Blackwell Ltd.
- CALLEAR, D. (1994) *Prolog programming for students*. London, DP Publications.
- CHANDLER, I. (1987) *Building technology. Design, production and maintenance*, London, The Mitchell Publishing Co. Ltd.
- CHANTER, B. and SWALLOW, P. (1996) *Building maintenance management*. Oxford, Blackwell Science Ltd.
- CHARTERED INSTITUTE OF BUILDING (1984) *Managing building maintenance*. HARLOW, P. ed. Ascot, The Institute.
- CHOMSKY, N. (1962) Explanatory models in linguistics. In: NAGEL, E. et. al., eds. *Logic, methodology and philosophy of science*, Stamford, Stamford University Press, pp. 518-550
- COLLINS, A. M. and LOFTUS, E. F. (1975) A spreading activation theory of Semantic Processing. *Psychological review*, 82, (No. 6), pp. 407-428
- DEPARTMENT OF THE ENVIRONMENT. National Consultative Council Standing Committee. (1979) *Project information: it's content and arrangement; a report and proposals on the way forward*, London, HMSO.

- DANLOS, L. (1987) *The linguistic basis of text generation*. Cambridge, Cambridge University Press.
- DOVBENKO, M. A. and UMANSKI, A. N. (1980) A comparative analysis of information retrieval thesauri...In: MIKHAILOV, A. I. ed. *International Forum on Information and Documentation*, 5, (4), Moscow, Vsesoyuznyi Institut Nauchno-Tekhnicheskoi Informatsii, pp. 9-11
- EDEN, C. and HARRIS, J. (1975) *Management decisions and decision analysis*. London, Macmillan.
- EDGILL, B. C. (1978) Relevance of information syntax to computer data structure. In: MONETA, J., ed. *Information technology*, Amsterdam, North-Holland Pub. Co., pp. 753-758
- ELLIS, G. (1908) *Modern practical joinery*. 3rd. ed., (facsimile pub. 1987), London, Stobart and Son Ltd.
- FIELD, B. J. (1975) *Semi-automatic development of thesauri using free-language vocabulary analysis*; R75/24, London, Institution of Electrical Engineers.
- FLEMING, M. C. and NELLIS, J. G. (1991) *The essence of statistics for business*. Hemel Hempstead, Prentice & Hall.
- FLETCHER, L. and MOORE, T. (1979) *Standard phraseology for builders quantities*. (4th ed) London, Godwin.
- FOSTER, H. P. et al. (1969) A study of coding and data co-ordination for the construction industry: Mechanical engineering services. *Current Paper*, 6/69, (March), Watford, Building Research Station.
- GARMAN, M. (1990) *Psycholinguistics*. Cambridge, Cambridge University Press.
- GATES, R. E. (1972) Comparative estimating for indirect work evaluation. *Work study & management services*, (Feb.), pp. 85-9
- GAZDAR, G. and MELLISH, C. (1989) *Natural language processing in PROLOG*. Wokingham, Addison-Wesley Publishing Company.
- GLEITMAN, L. R. (1965) Coordinating conjunctions in English. *Language*, 41, pp. 260-293
- GILLUM, T. L. (1964) Compiling a technical thesaurus. *Journal of chemical documentation* 4, pp. 29-32.
- GOODWIN, M. (1991) *Power of...QBasic*. New York, MIS Press.
- HALFF, H. M. et al. (1976) A context-sensitive representation of word meanings. *Memory and cognition*, 4, (pt. 4), pp. 378-383

- HARDIE, R. G. (1992) *Collins pocket English grammar*. London, Harper Collins.
- HARRIS, Z. (1957) Co-occurrence and transformation in linguistic structure. *Language* 33, pp. 288-340
- HASTINGS, J. (1969) Schedule non-repetitive jobs with the slotting technique. *Industrial engineering*, (April), pp.47-51
- HERBSMAN, Z. and WALL, R. A. (1987) An expert system for estimating construction cost. In: ISRAEL INSTITUTE OF TECHNOLOGY: FACULTY OF CIVIL ENGINEERING, 2, *Proceedings of the 4th International Symposium on Robotics and Artificial Intelligence in Building Construction*, June, Haifa, Israel Institute of Technology, pp. 208-221
- HILL, T. (1991) *Production/Operations management*. 2nd. ed. London, Prentice Hall.
- HOLM, E. E. and RASMUSSEN, L. E. (1961) Development of a technical thesaurus. *American documentation*, (July), pp. 184-190.
- HUDSON, R. A. (1976) *Arguments for a non-transformational grammar*, Chicago, University of Chicago Press.
- JAMES, M. (1992) *QBasic, the language of MS-DOS*. Leyburn, I/O Press.
- JOHNSON-LAIRD, P. N. and WASON, P. C. eds. (1977) *Thinking*. Cambridge, Cambridge University Press.
- JOHNSTON, H. P. (1969) A study of coding and data co-ordination for the construction industry: Electrical design and contracting. *Current paper*, 7/69, (March), Watford, Building Research Station.
- JONES, K. P., (1981) Problems associated with the use of compound words in thesauri, with special reference to BS 5723:1979. *The journal of documentation*, (June), 37, (2), pp. 53-68
- KAMIN, J. et. al. (1993) *Using MS-DOS 6*. Special ed. Carmel Ind., Lloyd J. Short.
- LEACH, E. (1976) *Culture and communication*. Cambridge, Cambridge University Press.
- LEE, A. M. (1966) *Applied queuing theory*. London, Macmillan.
- LEECH, G., (1989) *An A-Z of English grammar and usage*. Walton-on-Thames, Thos. Nelson and Sons Ltd.
- LIYOU, Y. I. et al. (1990) Methodology for knowledge acquisition in a group decision support system environment. *Knowledge acquisition*, 12, (Pt. 2), pp. 129-144
- LOVEDAY, R. (1969) *Statistics*. 2nd. ed., Cambridge, Cambridge University Press.

- MacKAY, D. M. (1961) The informational analysis of questions and commands. In: CHERRY, C. ed. *Information theory*, London, Butterworths.
- McCAFFER, R. et. al. (1991) The future of IT - an academic viewpoint, HARLOW, P. ed. *Technical Information Service*, No. 132, Ascot, Chartered Institute of Building.
- MARTIN, J. L. N. (1990) *BMI Special report: measured term contracts*. Serial 193, London, Building Maintenance Information Ltd., pp. 19-26
- MILLER, G. A. (1956) The magical number seven, plus or minus two: some limits on our capacity for processing information, *The psychological review*, (March), 63, No. 2, pp. 81-97.
- \_\_\_\_\_. (1965) Some preliminaries to psycholinguistics. *The American psychologist*, 20, pp. 15-20
- \_\_\_\_\_. and SELFRIDGE, J. A. (1970) Verbal context and the recall of meaningful material. In: POSTMAN, L. and KEPPEL, G. eds. *Verbal learning and memory*, Harmondsworth, Penguin Books.
- MUNDLE, C. W. K. (1979) *A critique of linguistic philosophy*. 2nd. ed., London, Glover & Blair Ltd.
- NEEDHAM, R. M. and SPARCK JONES, K. (1964) Keywords and clumps. *Journal of documentation*, 20, No. 1, (March), pp. 5-15
- NELSON, J. I' a, (1969) A study in coding and data co-ordination for the construction industry: Instructions to operatives. *Current paper*, 10/69, (March), Watford, Building Research Station.
- NEWTON, S. (1987) Dealing with dialectics. In: ISRAEL INSTITUTE OF TECHNOLOGY: FACULTY OF CIVIL ENGINEERING, 2, *Proceedings of the 4th International Symposium on Robotics and Artificial Intelligence in Building Construction*, (June), Haifa, Israel Institute of Technology, pp. 53-68
- O'BRIEN, J. and COMPTON, R. (1987) Field studies of the cognitive processes involved in construction activity. In: ISRAEL INSTITUTE OF TECHNOLOGY: FACULTY OF CIVIL ENGINEERING, 1, *Proceedings of the 4th International Symposium on Robotics and Artificial Intelligence in Building Construction*, (June), Haifa, Israel Institute of Technology, pp. 322-339
- ORTHOLOGICAL INSTITUTE, (1960) *The general Basic English dictionary*. London, Evans Brothers Ltd.
- PACKER, A. D. (1996) *Building measurement*. Harlow, Longman.
- PROLOG DEVELOPMENT CENTER, (1990) *PDC Prolog version 3.2: User's guide*, Copenhagen, Prolog Development Center.

\_\_\_\_\_ (1990) *PDC Prolog version 3.2: Reference guide*, Copenhagen, Prolog Development Center.

RADFORD, A. (1988) *Transformational grammar*. Cambridge, Cambridge University Press.

RANDALL, R. H. (1978) Building defects and feedback in PSA. *Maintenance Information Service*, No. 4, (Summer), Ascot, Institute of Building,

SAWYER, W. W. (1955) *Prelude to mathematics*. Harmondsworth, Penguin.

SCHANK, R. and ABELSON R. (1977) *Scripts, plans, goals and understanding*. New Jersey, Laurence Erlbaum Associates.

SHER, W., (1995) Classification and coding of data for computer-aided estimating systems. HARLOW, P. ed. *Construction papers*, No. 55, Ascot, Chartered Institute of Building.

SKIDMORE, S. and WROE, B. (1988) *Introducing systems analysis*. Manchester, The National Computing Centre Ltd.

SNOOK, K. (1995) CPI- Co-ordinated project information. HARLOW, P. ed.. *Construction papers*, No. 43, Ascot, Chartered Institute of Building

SPON'S CONTRACTORS' HANDBOOK (1991) *Minor works, alterations, repairs and maintenance*. 4th. ed. SPAIN & Ptnrs, ed. London, E. & F. N. Spon.

STERLING, L. and SHAPIRO. E. (1986) *The art of Prolog*. Cambridge, Mass. and London, The MIT Press.

THIERAUF, R. J. and KLEKAMP, R. C. (1975) *Decision making through Operations Research*. 2nd. ed. New York, Wiley International.

THOMAS, H. R. and SMITH, G. R. (1987) Conceptual design of an automated real-time data collection system for labor-intensive construction activities, In: ISRAEL INSTITUTE OF TECHNOLOGY; FACULTY OF CIVIL ENGINEERING, 2, *Proceedings of the 4th International Symposium on Robotics and Artificial Intelligence in Building Construction*, (June), Haifa, Israel Institute of Technology, pp. 849-862

TURNBULL, J. (1987) A guide to commissioning maintenance manuals. HARLOW, P. ed. *Technical Information Service*, No. 83, Ascot, Chartered Institute of Building.

VASCONCELLOS, M. ed., (1988) Technology as translation strategy, *American Translators Association. Scholarly monograph series*, II, New York, State University of New York.

WHEELER, A.J. and CLARK, A. V. (1992) *Building quantities: worked examples*. Oxford, Newnes.

WILD, R. (1989) *Production and operations management*. 4th. ed., London, Cassell Education Ltd.

WOLFF, J. G., (1977) Recoding of natural language for economy of transmission or storage. *The computer journal*, 21, (pt. 1), pp. 42-44