UNIVERSIDADE FEDERAL DE SANTA CATARINA

PÓS-GRADUAÇÃO EM LETRAS/INGLÊS E LITERATURA

CORRESPONDENTE

LIMITATIONS AND POSSIBILITIES OF MACHINE TRANSLATION:

A CASE STUDY

Por

ACÁCIA ROSAR

Dissertação submetida à Universidade Federal de Santa Catarina em cumprimento

parcial dos requisitos para obtenção do grau de
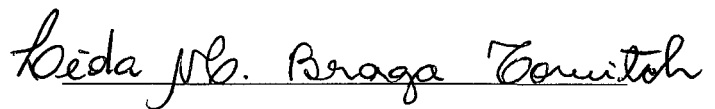
MESTRE EM LETRAS

FLORIANÓPOLIS

Junho 2001

Esta Dissertação de Acácia Rosar, intitulada Limitations and Possibilities of Machine Translation: a Case Study, foi julgada adequada e aprovada em sua forma final, pelo Programa de Pós-Graduação em Letras/Inglês e Literatura Correspondente, da Universidade Federal de Santa Catarina, para fins de obtenção do grau de
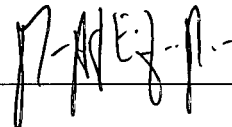
MESTRE EM LETRAS

Área de concentração: Inglês e Literatura Correspondente
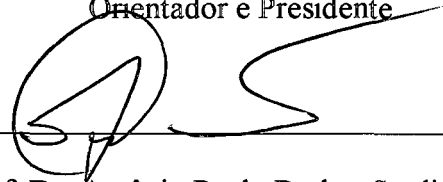Opção: Língua Inglesa e Lingüística Aplicada

Prof. Dr. Lêda Maria Braga Tomitch
Coordenadora

BANCA EXAMINADORA:

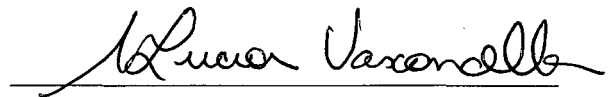Prof. Dr. Marco Antônio Esteves da Rocha
Orientador e Presidente

Prof. Dr. Antônio Paulo Berber Sardinha
Examinador

Prof. Dr. Maria Lúcia Vasconcelos
Examinadora

Florianópolis, 07 de Junho de 2001

ACKNOLEDGEMENTS

All things work for good to them that love God... Rm. 8:28 (Bible)

Let your light shine before men, that they may see your good deeds and praise your Father in heaven. Mt. 5:16 (Bible)

ABSTRACT

LIMITATIONS AND POSSIBILITIES OF MACHINE TRANSLATION:

A CASE STUDY

ACÁCIA ROSAR

UNIVERSIDADE FEDERAL DE SANTA CATARINA

2001

Supervising Professor: Dr. Marco Antonio Esteves da Rocha

This thesis presents results of a case study on renderings of the English pronoun *it* into Portuguese, following a brief up-to-date overview of the field since its beginnings up to nowadays. A roughly forty-five thousand-word parallel corpus of written source and target language text was collected. An annotation specifically designed for the purposes of this study was used to classify 305 tokens of *it*. Elements included in the annotation scheme are syntactic function, antecedent type and processing strategy, which are discussed in the thesis. Results are then compared to renderings of two commercial machine translation systems, having solutions by human translators in the corpus as a parameter. Suggestions are made concerning possible improvements of existing systems on the basis of corpus evidence. Some aspects of corpus evidence are weighed against principles of present-day approaches to machine translation, in an attempt to enrich the discussion on recent tendencies in the field.

Number of pages: 99
Number of words: 30.759

RESUMO

LIMITAÇÕES E POSSIBILIDADES DA TRADUÇÃO DE MÁQUINA:

UM ESTUDO DE CASO

ACÁCIA ROSAR

UNIVERSIDADE FEDERAL DE SANTA CATARINA

2001

Professor Orientador: Dr. Marco Antonio Esteves da Rocha

Este trabalho apresenta resultados de um estudo de caso sobre a tradução do pronome inglês *it* para o português. Apresenta também um breve panorama geral do desenvolvimento da tradução de máquina desde seu início até a atualidade. Um *corpus* paralelo de aproximadamente quarenta e cinco mil palavras das línguas de partida e chegada foi coletado. Também foi utilizado um esquema de anotação especificamente desenvolvido para os propósitos deste estudo, a fim de classificar as 305 ocorrências do pronome *it*. Os elementos que compõem a anotação são: função sintática, tipo de antecedente e estratégia de processamento, os quais são discutidos nesta dissertação. Os resultados são comparados a traduções de sistemas comerciais de tradução de máquina, tendo como parâmetro soluções apresentadas por tradutores humanos no *corpus*. Sugestões são feitas quanto a possíveis melhorias dos sistemas existentes com base em *corpus*. Alguns aspectos da abordagem de *corpus* são comparados com os princípios das presentes abordagens de tradução de máquina, numa tentativa de enriquecer a discussão sobre as atuais tendências nesta área.

TABLE OF CONTENTS

LIST OF TABLES AND GRAPHS

ABBREVIATIONS

SL	Source Language
TL	Target Language
SLT	Source Language Text
TLT	Target Language Text
ALPAC	Automated Language Processing Advisory Committee
MT	Machine Translation
RBMT	Rule-Based Machine Translation
EBMT	Example-Based Machine Translation
BNC	British National Corpus
LFG	Lexical Functional Grammar
HT	Human Translation
PTP	Power Translator Pro

CHAPTER 1

INTRODUCTION

> If I render word for word, the result will sound uncouth, and if compelled by necessity I alter anything in the order or wording, I shall seem to have departed from the function of a translator. Cicero (106-43 B.C.) *De optimo genere oratorum* v.14, *apud* Nord, 1997

Nowadays, the rapid development of computational means and the idea of globalization have spread all over the world. There is a need for communicating properly in any language. When people are faced with a foreign-language barrier, the usual way to deal with this situation is to find someone to translate or try to use technology to help them. It is within this context that Machine Translation (MT) is inserted.

It is perceivable, even for those who are not involved in the research field, that Machine Translation is still under development and not perfect in performing its task accurately. A visible example of this imperfectness is possible to find on Internet. Some people are using this kind of translation in their websites or to send e-mails to other nations. As a result, several groups and associations all over the world have been researching on this subject. One specific problem associated to MT regards anaphoric relations. This subject has called researchers' attention in recent years and as a result there is a vast literature about it.

This thesis discusses the way two commercial MT systems (*Power Translator Pro* and *Systran*) translated texts from different genres containing cases of anaphora, which, in the present research, will be restricted to tokens of the English pronoun *it*. It

also concentrates in comparing translation similarities of the pronoun in human made and machine translations. The questions which motivated the present study are:

- Were these programs developed within a corpus-based approach?

- What are the main strategies used to translate the pronoun "*it*"?

- How does the machine deal with the question of ambiguity?

- How could existing shortcomings be improved in order to make MT systems operational and efficient?

Unfortunately it was not possible to answer some of these questions due to lack of enough information provided by software manufacturers. My initial assumption is that the translation programs were not developed within a corpus-based approach and ambiguities are also not dealt with a special care thus resulting in non efficient systems.

As a whole, this thesis is arranged in a general-to-specific fashion. It is divided into six chapters, in an attempt to present how the English pronoun *it* was examined in terms of MT. In this chapter, I have briefly introduced my study, presenting its purpose and how it was organized.

Chapter 2 presents the general theoretical perspectives for this thesis. The first part provides a brief up-to-date overview of the MT field since its beginnings up to nowadays, while the other ones present a short discussion of the Rule-Based Machine Translation (RBMT) and Example-Based Machine Translation (EBMT) approaches in their main aspects, as well as theories concerning anaphora resolution.

Chapter 3 brings up the methodology used in the research informing how data were collected, the analytical procedures adopted, their evaluation and statistics.

Chapter 4 examines and tests how two existing translation software, *Systran* and *Power Translator Pro* translate the pronoun *it* into Portuguese in different contexts, as well as the diagnosis of the translation limitations of each system. Still in Chapter four,

there is a discussion of statistical analysis for the results and suggestions for improvements in the software.

Chapter 5 presents a statistical analysis of corpus data and a thorough discussion of a processing strategy type, namely, collocation[1].

Finally, Chapter 6 draws the thesis to a close with a conclusion reporting the general findings of my investigation and making suggestions for further research.

CHAPTER 2

GENERAL THEORETICAL FRAMEWORK

> Texts are intended to be meaningful to their addressees, and therefore, text producers (and translators are text producers, too) shape their texts so as to conform as far as possible to the situational conditions of the addressees. Nord, 1997

The rapid development of Translation Studies in Europe during the 1950's was a result of the emphasis on linguistics and experiments with machine translation. Translation Studies, as a discipline, has been developed on the basis of practical applications and at present covers a wide field of research. In order to understand and systematize its object of study, Bassnett (1994) proposes a division into four general areas of interest, each with a degree of overlap. Two of these areas are *product-oriented*, which means that the functional aspects of the target language text (TLT) in relation to the source language text (SLT) are stressed. On the other hand, the other two are *process-oriented*, i.e. the stress is on the analysis of what actually takes place during translation.

The first division of Translation Studies proposed is *History of Translation*, which is seen as part of literary history. The history of translation also includes the investigation of translation theories at different periods, the critical responses to translation, the role and function of translations in a given moment, the methodological development of translation, as well as the analysis of the work of individual translators.

The second division, *Translation in the Target Language (TL) Culture*, focuses on the influence of individual texts, authors or genres over the TL system, as well as the

absorption of the norms of the translated text into the TL system and the principles of selection operating within that system.

As to the third division, *Translation and Linguistics*, it emphasizes studies on the comparative arrangement of linguistic elements between SL and TL text with regard to phonemic, morphemic, lexical, phrasal and syntactic levels. Moreover, this branch includes the study of the problems of linguistic equivalence, language-bound meaning, linguistic untranslatability, machine translation and also the studies of translation problems of non-literary texts.

Finally, the fourth division, *Translation and Poetics*, embraces the whole area of literary translation, in theory and practice. Studies in this area may be general or genre-specific, including investigation of particular problems of translating poetry, drama and subtitling.

As mentioned before, Machine Translation belongs to the third category above and is defined in the *Dictionary of Translation Studies* (1997) as:

> a term used to refer to translation which is performed wholly or partly by computer. As implied by the definition, such translation may be carried out with or without human intervention; however, if there is a considerable level of intervention, or if computer applications are simply used as "translation tool", then it is more common to talk of Machine-Aided Translation. (p.99)

2.1 – Background

The idea of using MT to provide translations from one language into another appeared in the 30's but it was not after World War II that actual systems started being developed. During the 50's, several groups began to do research on machine translation systems known as direct systems (direct approach) but there were few positive, encouraging results. These systems just acted as automatic bilingual dictionaries

because they simply replaced "Source Language (SL) words with Target Language (TL) equivalents and changed the order of words" (Sampson, 1987, p.91). The machines at the time were far less powerful and much more difficult to handle than the ones used nowadays. This made the sort of complex processing required for MT very difficult indeed. The systems were very limited in relation to the kind of data they could handle and required so much human post-editing that they proved to be more expensive than a human translator. Another relevant problem at that time was the "lack of a sufficiently sophisticated linguistic theory to provide a frame of reference for the task that MT needed to undertake" (Crystal, 1998, p. 352).

Concerning the results achieved by these initial efforts, the Automatic Language Processing Advisory Committee (ALPAC) summarized in a 1966 report all their dissatisfaction. They concluded that "human translating was faster, more accurate, and less expensive than MT, and that no further support for the latter should be provided" (Crystal, 1998, p.352). The impact of the ALPAC report on existing MT research was very strong, leading to a severe reduction in the availability of funds and to a general feeling of mistrust in such research initiatives amid the scientific community. Thus, only a small amount of MT research was carried on in subsequent years.

It was only in the 70's, with the development of linguistics and computer science (particularly research programs in artificial intelligence), that machine translation got a new promising perspective. The main developments were to create systems of analysis that could deal with grammatical and semantic complexity. Machine translation researchers, based on Chomskyan deep structures[2], knew that people could not expect to translate a text appropriately unless they understood it. In other words, understanding a text before translating it is a crucial task. The notion of deep structures were the basis for the attempts to codify this understanding into syntactic and semantic rules of the

generative kind, typically using context-free grammars (the S ← SN SV structure). This was the dominant paradigm of research at the time. The paradigm was also adequate to a technological reality in which the cost of memory still made corpus-based approaches too expensive computationally. Nonetheless, it was during this time that second generation systems appeared. They embodied the interlingua idea that the representation meant to get the complete meaning of the text and to be available equally well in both source and target languages. According to Whitelock & Kilby (1995), the interlingua means that:

> Translation is a two-stage process. A structure is assigned to a SL text, using knowledge of the source language only. This structure is a sentence in a universal language, representing the 'meaning' of the text. From it can be generated an equivalent utterance in any TL, without regard for the original SL. (p. 6)

Later on, a third generation of systems emerged, emphasizing transfer. These systems contrasted the interlingual idea that translation involved filling in the gaps in meaning in a universal language, discussed in Whitelock & Kilby (1995) with the meaning-situation view. In the "filling gaps" discussion it is assumed that translation involves the use not only of conscious but also unconscious knowledge. In other words, the translator should use his/her knowledge and own strategies to translate and find the interpretation to fill the gaps of information not explicit in the text. On the other hand, the idea of "meaning" in a sentence varies according to the circumstances in which this sentence occurs. This happens because when we talk about a linguistic form, there is a large variety of possible meanings to it. It is necessary to define and limit the information/"meaning" to be computed into a MT system so as to enable it to function more efficiently.

In transfer, the system designer is primarily an expert in the TL. He/She is responsible for analyzing and deducing the questions that will be asked when the

computer is translating the SL text. In fact, he/she is supposed to provide the answers to all the possible questions that may appear in the system. In addition to that, Whitelock et al.(1995) say that an "MT system should incorporate an initial monolingual analysis of the SL text, followed by a phase using bilingual knowledge" (p.9). They also state that monolingual synthesis in MT transfer has little importance. If multiple translation equivalents cannot be sorted out in order to select a single best choice, on the basis of bilingual knowledge, an arbitrary choice will be made by synthesis. This can be different if a more sophisticated treatment of stylistic factors is included into the system. This synthesis should not be considered the opposite of analysis but something deterministic.

It is remarkable that even though the systems mentioned above are quite different from one another, interlingua and transfer have something in common, namely the use of rules. This rule-based MT approach will be discussed in the section 2.3.

2.2 – A New Perspective

In contrast to all MT systems mentioned before, built according to an approach broadly known as Rule-Based Machine Translation (RBMT), the 80's witnessed the steadily growing importance of systems which attempted to achieve automatic translation by using a large corpus of examples of translation, where an example is considered as a pair composed of a source text and its translation. The approach used by these systems is known as Example-Based Machine Translation (EBMT), and has been developed along with corpus-based approaches to natural language processing in general. Jones (1996) states that EBMT is different from RBMT in that "it seeks to avoid the well-known difficulties and complexities of purely compositional translation

by treating translation equivalents as larger units of pre-composed text" (p.14). Although, it may be said that EBMT clearly seems to be an improvement, this does not mean that MT problems are to be easily solved once EBMT is adopted because language is original and innovative and, as a consequence, it would probably be almost not viable to get all possible examples. Besides that,

> a very large amount of duplication would begin to occur in the
> expression of examples, with the added disadvantage that there would
> be no ability to give a generalized account of how language works.
> (Jones, 1996, p.14)

Having the existing shortcomings discussed above in mind, a number of research initiatives have focused on specific 'hard' problems, which have systematically hindered the development of accurate MT systems. Complex textual links, such as anaphoric relations[3], are widely acknowledged as one of these difficulties.

This difficulty is not different in what regards the rendering of English texts into Portuguese and vice-versa. Whereas the English referring system relies on explicitly stated pronouns, such as *it*, to achieve textual cohesion, Portuguese allows the omission of subject pronouns and their treatment as non-existing for certain special uses of very common verbs like "ser (be)", "estar (be)" and "haver (there to be)". A very large amount of research has been carried out concerning anaphoric relations, both within the scope of linguistic theory and in studies more directly linked to psycholinguistics[4] and artificial intelligence[5]. Among the latter, not many have focused on specific MT problems such as the rendition of English texts where the pronoun *it* occurs into Portuguese.

Accurate translations involving resolution to the English pronoun *it*, which is a common difficulty in MT, also requires knowing what the pronoun actually refers to, if one has to translate the pronoun into a gender-marked language such as Portuguese. The

appropriate pronouns have thus to be selected. In addition, language is often ambiguous, and MT systems have to make arbitrary choices when handling ambiguous passages because ambiguity will not usually be translated into a TL. Another point that should be carefully dealt with in translation refers to the matter of designing a computer program. Sampson (1987) argues that:

> the range of possible answers to a programming problem is a perfectly closed system - it is defined by the syntax of the programming language. And all the elements of the programming language are perfectly rigorously specified, so if you know what your program is supposed to be doing then in principle purely deductive reasoning ought to give you a program which functions correctly. (p.104)

This leads to a conflict between the typically ambiguous nature of textual relations, such as anaphora, and programming languages. Thus, applications such as MT have to find ways to deal successfully with this conflict.

2.3 – Rule-Based Machine Translation – RBMT (Interlingua/ Transfer)

As it has been mentioned in the preceding section, the first MT systems were known as Rule-Based ones. The Rule-Based systems were part of three important generations known as direct, interlingual and transfer approaches.

Whitelock & Kilby (1995) discuss the historical development of such systems from first-generation (direct) systems to third-generation (transfer) systems (that incorporated a monolingual analysis of the SL text followed by a phase using bilingual knowledge). Through analysis, they perceived that MT is appropriate for a limited type of texts called "nonliterary" or "informative". These texts are basically technical books, manuals, etc. They also noted some linguistic problems that occur while translating from one language into another. They proposed to divide these problems into two

different categories; problems for analysis (that can be said to display various types of ambiguity) and problems for transfer. Among the problems included in the first category mentioned above, the problem of ambiguities created by the need to identify the correct antecedents of anaphors is particularly relevant to the present work, as it will be made clear later on.

Whitelock & Kilby also present a brief description of the most well known machine translation systems, especially *Systran*. *Systran* is probably the oldest and best-known existing system and also one of the most widely used to this date, with a reasonable degree of success. There are those who say that such success is the result of incorporating linguistic knowledge into its procedures.

In contrast to Whitelock & Kilby's classical description of RBMT, Van Eynde (1993), suggests an interaction between monolingual and bilingual knowledge in these systems. The division of labor between these two modules in MT is a bone of contention in the transfer versus interlingua discussion, and it is approached from the point of view of linguistic motivation. According to Van Eynde, the result of the combination of the two forms of knowledge mentioned above is a proposal for constructing language independent representations in terms of which language specific grammatical elements can be analyzed and model-theoretically interpreted:

> Since the method is only meant to apply to grammatical elements, the resulting MT systems will be mixed: the lexical elements will be treated in a language pair specific way (transfer), whereas the grammatical elements will be treated in a language pair independent way (interlingua). (p.40)

Concerning the transfer model, which involves stating correspondences between SL and TL representations, in her article *Co-description and Translation*, Sadler (1993) discusses and evaluates a proposal for translation within the Lexical Functional Grammar (LFG)[6] formalism, with reference to both the transfer model and the proposals

to use unification-based grammars for translation. There is a detailed introduction to Kaplan & Netter's (1989) theory to MT based on co-description, using the equality and description based mechanisms of LFG, which extends the discussion about the behavior and adequacy of the LFG approach in relation to some cases of 'difficult' transfer. One of these difficulties is related to cases in which languages do not share the same head-dependent relations. Sadler argues that such cases are significant in representing a great structural difference involving languages which a translation formalism should be able to deal with in a natural way. Moreover, she states that: "representation-based transfer systems which apply rules to recursively translate source language structures would seem to avoid this problem in principle" (p.65).

2.4 - Example-Based Machine Translation – EBMT

It was only in the eighties that researchers started noticing the importance of using a large corpus of examples to achieve automatic translation. This large linguistic corpus is the defining feature of an approach to linguistics now widely known as corpus linguistics. It is defined by Francis (1992) as a "collection of texts assumed to be representative of a given language, dialect, or other subset of a language, to be used for linguistic analysis" (p.17). In other words, corpus linguistics in machine translations can be understood as a corpus which use a set of examples from which the software analogizes when producing a translation.

Having this same view, Sampson (1987) presents a new perspective to translation and the translator work using some authentic-language examples from his own research. He discusses, through a number of occurrences (extracted from the Lancaster - Oslo- Bergen Corpus), the problem of solving the reference of the English

pronoun *it*. He states that if someone wants to translate this pronoun into a gender-marked language, it is important to know what it is referring to. Sampson also criticizes the computational linguistics community attitude of picking from Chomskyans, the biased technique of working only with small and carefully-chosen samples of language. Books emanating from Chomskyan School have usually created example sentences with selected problems to which answers are already known. Sampson argues that successful systems, which have a practical application, should use real-life language, extracted from naturally occurring corpora, as a basis for their approach to MT:

> My belief is that systems which process real-life language successfully will not be decomposable into a core system that deals with a dapper, polished little subset of language, together with peripheral components to handle 'performance deviations'; successful systems will be geared to real-life language through and through. (p.106)

Besides Sampson's discussion about EBMT, a program for aligning sentences in bilingual corpora was developed in the nineties. Aligning a sentence means to identify similarities between sentences in one language and sentences in the other language. According to Gale & Church (1993), aligning sentences is the first step toward constructing a probabilistic dictionary for use in aligning words in Machine Translation, or for constructing a bilingual concordance for use in lexicography. They also describe a method and a program (align) for aligning sentences based on a statistical model of sentence length. Such a program takes into consideration the fact that long sentences in one language tend to be translated into long sentences in the other language, the same thing occurring to short ones. There is a probabilistic score to each pair of proposed sentence pairs. This score is based on the proportion of lengths of the two sentences and the variance of this proportion. The probabilistic score is used to find the greatest likelihood of sentence alignment types.

Gale & Church's idea is significant to this present work because, once the aligning process is well developed, it will be extremely useful to help identifying specific problems in translating the pronoun *it* from one language into another.

2.5 - Anaphora Resolution

Anaphora resolution is of great importance when discussing MT in general and particularly in this thesis. According to Rocha (1997), anaphora is "the relationship between a term – called the anaphor – which must be linked to an explicit or inferable element in the discourse – called the antecedent – in order to successfully accomplish semantic interpretation." In other words, the anaphor cannot be correctly interpreted without the identification of the antecedent. In addition, Mitkov (1996) emphasizes that the establishment of the antecedent is extremely important especially when translating into languages in which the resolution of anaphoric relation depends mainly on the identification of gender marks. Furthermore, he states that the great majority of MT systems do not rely on this approach and, as a consequence, their results do not go beyond the sentence level. This means that not going beyond the sentence level leads to difficulties in assigning correct gender markings to pronouns in the TL. Mitkov also affirms that anaphora resolution as analysis is a hard difficult task;

> "but translation adds a further dimension in so far as the reference to a
> discourse entity encoded by a source language anaphor by the speaker
> (or writer) has not only to be identified by the hearer (translator or
> translation system) but also re-encoded in a co-referential/
> cospecificational expression of a different language" (p.7)

Pronouns are the most common type of anaphora. The English third person pronoun *it*, used as an anaphor, has been a subject of discussion among some computational linguists. Rocha (1998) proposes an annotation scheme (which will be

discussed in Chapter 3) to help determining the relationship between the anaphor and its antecedent which includes an assessment of the type knowledge (called *processing strategy*) used to process anaphoric references identified in the corpus. From the categories which compose this processing strategy, collocations account for an important share of the total number of cases analyzed.

The concept of collocation is well known among computational linguists in general. In his dictionary, Harris (1995) defines it as "the habitual association or co-occurrences of words within some grammatical structure, as *submit a proposal, pen and ink.*" Hoey (1991) also states that "collocation has been the name given to the relationship a lexical item has with items that appear with greater than random probability in its (textual) context" (p.7). This relationship can be generally demonstrable through statistics. He also argues through an evaluation on Phillips's (1985) work, that collocation is "text-sensitive and that clusters of repetitions occur irregularly (and therefore significantly) at long distance" (p.22).

In addition to that, Sinclair (1991) states that the role of collocation is to illustrate the idiom principle (which means that "a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analyzable into segments" p110). On some occasions, words appear to be chosen in pairs or groups and these are not necessarily adjacent. Sinclair also emphasizes that one important aspect of collocation is when two words of diverse frequencies collocate significantly, the collocation having a different importance in the description of each of the two words. For instance, when a word that is being studied appears followed by a less frequent one, this is called *downward collocation*. When this process is inverted it is called *upward collocation*. "Upward collocation, of course, is the weaker pattern in statistical terms, and the words tend to be element of

grammatical frames, or superordinates. By contrast, downward collocation gives us a semantic analysis of a word" (p.116).

Carreiras & Oakhill (1993) discuss all the foregoing topics, emphasizing that a theory of discourse comprehension must have, as a subpart, a theory of the interpretation of anaphoric expressions. By means of a case study it was possible for them to say that "in English, pronouns are marked for gender and number and, at least under certain circumstances, pronouns that can be resolved solely on the basis of these markings are understood more quickly than those that cannot" (p.98).

However, they also argue that the "agreement between a pronoun and its antecedent is either purely semantic, or if there is agreement at a more superficial level, it goes hand in hand with semantic agreement" (p.98). For instance, if someone is introduced using a female or male proper name, he/she can be later on referred to by the equivalent third person pronoun. This pronoun will match perfectly in terms of gender and sex of the person mentioned.

In short, both RBMT and EBMT approaches are significant to understanding how the first MT systems appeared and how they were developed until nowadays. Anaphoric relations have attracted the attention of researchers in the last years. Their resolution is important because translating the pronoun *it* correctly into Portuguese is a difficult task for the machine to perform since Portuguese is a gender language. Moreover, anaphoric relations should be studied more in order to improve current systems.

The next Chapter presents the methodology employed in the research as well as a more detailed description of the annotation scheme with real-life language examples extracted from the corpus. As it will be noted later on, the annotation is also significant

to the research because it can classify all anaphora cases thus preventing me from not analyzing the most problematic ones.

CHAPTER 3

METHODOLOGY

> For many years the MT world has been quite a small one, with few research programmes and commercial organizations involved. This situation is now dramatically changing.
> Crystal, 1998

3.1 – Data Collection

The use of a corpus in this work was the central aspect of my research effort towards an investigation of anaphoric phenomena for the purposes of improving MT software. So the ideal solution was to collect a parallel corpus of English originals and renderings of these originals into Portuguese by professional translators. The corpus was all collected through a scanning and data storage process. This approach reproduces the typical EBMT approach to a translation problem. According to Shuttleworth (1997), the advantage of using a parallel corpus is that it provides information on patterns of:

> "specific target texts, and so give insight into the particular translation practices and procedures which have been used by the translator. Full exploitation of the potential of parallel corpus is possible only with the use of special software to enable the investigator to align ST sentences with their TT equivalents or to conduct bilingual concordancing operations". (p.120)

Through this parallel corpus analysis it was believed useful resolution patterns would emerge to support suggestions for improvements in MT software, following the annotation of all anaphora cases and by comparing MT software solutions to human solutions and to annotated tokens of it in the English originals.

The data analyzed in this study are composed of 305 occurrences of the English pronoun *it* extracted from ten texts of different genres. The purpose was to collect a significant number of anaphora cases for subsequent testing and statistical analysis. The occurrences were all extracted with the help of the software *Wordsmith Tools* (developed by Mike Scott and published by Oxford University Press). The computer translation of the occurrences, which is part of the Portuguese corpus, was generated by two different MT systems: *Systran* and *Power Translator Pro* (*PTP*). All the English texts obtained from professional translators residing in Florianópolis included a Portuguese translation of the originals produced by the translators themselves, so as to provide independent parameters of comparison and evaluation and avoid being influenced by the MT results. The texts emcompass thematic areas such as Biology, Dentistry, Nursing, Tourism and Computing, and add up to a 45,866- word corpus distributed as follows:

- Artificial Sucking Habits: Etiology, Prevalence and Effect on Occlusion – 3.996.

- CLIA Cruising: The Best Vacation – 3.746.

- Marginal Fit of Gold Inlay Castings – 1.895.

- Network Management Strategy and Architecture – 4.167.

- Nursing as Comforting: The Artistry of Caring – 4.383.

- Periodontal-Response to Long-Term Abuse of the Gingival Attachment by Supracrestal Amalgam Restorations – 2.137.

- Physiologic Dimensions of the Periodontium Significant to the Restorative Dentist – 2.096.

- Pinnipedia: Seals, Sea Lions and Walruses – 6.487.

- Restorative Dentistry and the Periodontium – 3.843.

- SWIFT3D: Userguide – 13.116.

From all the texts above, *Nursing as Comforting: the Artistry of Caring*, is the only academic paper which has a narrative written by another person inserted it, a non-typical characteristic of academic writing.

It should be pointed out that some texts from the corpus were originally written in English by non-native speakers (from Israel and Sweden), while others I am not sure about their origins because writers are members of American university faculties but having foreign family names.

3.2 – Basic Methodology Employed in Writing

It is important to mention that American English is used in this thesis, even though some British words may be found "mainly" in citations of the writers' own words. The examples in English and Portuguese used in the thesis were extracted from the corpus. Other relevant aspects referring to matters of organization in the text and examples are the following ones:

- Caps lock when a term appears as an abbreviation.

- Italic lettering or simple quotation mark for emphasis or for designating foreign words.

3.3 – Statistical Analysis

After being extracted and annotated, the occurrences of the pronoun were all translated into Portuguese by two translation software. The results were all cross-tabulated, the *correct*, *incorrect* and *partially correct* categories[7] for automatic translation results against each of the properties in the annotation scheme, showing the

distribution of the *it* tokens according to their classification with approximate percentages. The use of statistical techniques to explore corpus data, together with the properties studied, was influential in shaping choices made for this thesis. All the statistical results presented in Chapters 4 and 5 were calculated with a chi-square test[8] in order to give a measure of how probable it was that a certain distribution of an independent variable across categories of a dependent variable was the result of chance. The statistical procedures mentioned above were all performed by the software *SPSS for Windows 8.0*, which is defined as a comprehensive system for analyzing data.

3.4 - Analytical Procedures

This section describes the three analytical procedures used as working routines during the research process and formed the main framework for corpus analysis. One is the standard used to determine when an MT result was to be considered good translation and the other is the annotation to classify *it* tokens in the SL corpus. The subsequent subsection begins with the latter and then the first procedure is described in the second subsection.

3.4.1 - Annotation Scheme

Once the occurrences were collected from ten different texts that composed the data, an annotation scheme was adapted from Rocha (1998) in order to classify the anaphoric relation in question. The annotation classifies all anaphora cases thus preventing the researcher from not analyzing the most problematic ones. If there are difficult cases that require language experience from the human analyst to solve them, I

can conclude that they are very difficult for a machine, which cannot rely on such experience and must be instructed in some way to deal with such situation. Through the classification it was possible to analyze each token of *it* according to three properties, namely: compared syntactic function, type of antecedent and processing strategy. The categories used to classify each token of *it* according to these properties are presented below.

The annotation scheme draws on work by Rocha (1998). The symbols used in the annotation are placed next to the category name, and they also appear in brackets after the tokens. Each category is presented with an example extracted from the corpus to contextualize and support the present research. The antecedent appears in italics and underlined while the occurrence of the pronoun *it* is in red. There are cases in which two antecedents from different tokens appear in the same sentence, so for the sake of organization one of them appears in blue color. The whole annotated corpus is provided in the Appendix 1.

3.4.1.1 – Compared Syntactic Function

This is a property that refers to the analysis of the syntactic function of each element in the anaphora relationship (that is, anaphor and antecedent), according to traditional syntactic functions within the English sentence. In other words, the syntactic function is assigned independently to each element and then compared. So, the pairing is classified based on the three categories below:

a) Same

The syntactic function is analyzed to identify if there is any parallelism between the antecedent and the anaphor. In other words, the compared syntactic function is

classified as **same** when antecedent and anaphor have the same syntactic function in their respective sentences. For instance:

(1) Goal: Relocate *your object* exactly where you want it [Same;Ex;FtC]

In the case above, both *your object* and the pronoun *it* are direct objects.

b) Different (Diff)

The category classifies cases in which anaphor and antecedent have different syntactic functions in their respective sentences.

(2) Write your Mom *a quick email.* It'll [Diff;Ex;FtC] make her happy. Plus, you've got a moment or two to kill.

c) Non-applicable (NA)

It occurs when the comparison of syntactic functions in their respective sentences does not apply to the analysis of the anaphor and the antecedent, hence "non-applicable". It may appear related to an explicit antecedent or, as in most cases, to an implicit or non-referential antecedent. The last two categories mentioned are very common in collocation cases and will be presented in Chapter 5.

(3) It [NA;Ex;Coll] has been claimed *that supracrestal injuries to the dentogingival attachment reversible provided the lesions are allowed to heal against biologically acceptable root surfaces.*

anaphora is "the relationship between a term – called the anaphor – which must be linked to an explicit or inferable element in the discourse – called the antecedent – in order to successfully accomplish semantic interpretation."

3.4.1.2 - Types of antecedent

This property refers to the identification of the antecedent by the analyst. It basically relates to the explicit/implicit dichotomy. However, a third category was

included, which classifies cases in which there is no antecedent. Although the distinction between an implicit antecedent and an explicit one is simple in most cases, it may involve fairly complicated decisions in some relatively rare occurrences. This property is divided into three categories:

a) Explicit (Ex)

The explicit referent is any identifiable element present in the text that precedes the anaphor. In some cataphoric[9] cases, the antecedent can be in the subsequent text. It may be a sentence, clause or chunk of discourse, provided it is clearly recognizable as realized in the text.

> (4) *Water* with its great resistance presents problems for things which move even slowly through it [Diff;Ex;DK].

b) Implicit (Im)

The implicit property occurs whenever the antecedent has not been previously introduced in the text nor relates cataphorically to the anaphor, thus it has to be inferred by changing the classification of the words according to the parts of the speech.

> (5) You can now treat the light as a full-fledged object, using the same mouse maneuvers you would use to position objects. The best way to aim a light placed into the scene is with the object Trackball. Although you can *resize * the light with the scaling tool*, it [Diff;Im;DK] has no effect on the amount of light cast. (Read: don't bother.) *resizing is the referent.

The example above shows that the implicit antecedent is another form of the main candidate[10], the most probable implicit element being is *resizing* since the antecedent belongs to a different part of speech related to *resize*, i.e., a noun deriving from a verb.

c) Non-referential (NR)

The non-referential property refers to the occurrences which are not really anaphoric, since there is no antecedent, i.e., a nonreferential pronoun. In my corpus, the most common cases are those related to time expressions. Rocha (1998) states that occurrences of anticipatory *it* should be classified as referential in all cases, that is, whether the occurrence is the subject of a cleft sentence or filled the position of an extraposed clausal constituent. Moreover, he also says that cases of obligatory extraposition, such as the anticipatory *it* as subject of the verb *seems*, are also considered as referential, as well as tokens of anticipatory *it* which refer to extraposed nonfinite clauses. It is important to mention that most of the tokens found in the corpus, which were classified as non-referential, appeared in contexts that were considered as collocations.[11]

(6) It [NA; NR; Coll] was after 6 o'clock when I slipped back into the room, just as the early morning light was coming through the "Mrs. P," I reached out and touched her arm. She raised tier tear streaked face to look at me. "It's [NA; NR; Coll] time," I said and waited When she was ready, I helped her off the, bed mid held her in my arms for a few moments. We cried together. "Thank you, nurse," she said as she look into my eyes and pressed my hand between hers. Then she turned and walked away.

3.4.1.3 - Processing Strategy

The inclusion of this property in the annotation scheme is an attempt to classify the cases of anaphora in relation to processing as an attempt to help guiding anaphora resolution in translation systems. It relates to the incorporation of psycholinguistic elements into the annotation. It might be argued that the two variables discussed in the other subsections are not sufficient to cover important processing information. Different strategies are needed to resolve the same type of anaphor, and frequently the same

anaphoric word or phrase in different contexts, as a result only a processing strategy can be helpful in these cases.

Four categories are used to classify cases according to this property:

a) First Candidate Search (FtC)

According to Rocha (1998), the strategy to identify the first candidate consists of a search for the antecedent on the basis of syntactic knowledge – basically the notion of command – and agreement constraints. The first suitable candidate found in such a search is the antecedent for the anaphor.

> (7) There is no right or wrong to see in this *situation* – it [Diff;Ex;FtC] is unique. As we listen, we might individually consider the calls for nursing we hear.

In the example above there is not any other possible candidate to be taken in consideration rather than the underlined one.

b) First Candidate Chain (FtCCh)

This kind of property has an anaphor as its antecedent. The last anaphor in the chain is linked to another and to a common antecedent, as it is shown in the following example:

> (8) Slowly drag towards the center of *the sphere* to make it [Diff;Ex;FtC] slightly smaller than it [Diff;Ex;FtCCh] was.

It is relevant to say that the anaphor which starts the chain is not necessarily resolved by a first-candidate strategyeven though it may in most occurrences.

c) Discourse Knowledge (DK)

This processing strategy refers to anaphora cases in which it is hard or complex to identify the antecedent and full processing of discourse is required for its resolution (bypassing a number of candidates to reach the correct antecedent).

> (9) *Blood* from the kidneys is drained into *the stellate plexus* which covers the kidneys. From here it [Same;Ex; DK] passes through one of 3 vessels to the branch of the posterior vena cava adjacent to it. [Same;Ex;DK]

In the first anaphora case in the example above, it is clearly identifiable that the antecedent is *blood*; however, the same does not occur in the second one. More than common processing is necessary as it involves lexical knowledge. If the translator does not know that *stellate plexus* means an entwine of many vessels it will be quite difficult to identify the referent because there are some possible candidates near the anaphor.

d) Collocations (Coll)

For the purposes of this research, a collocation is an identifiable combination of elements that contains a token of the pronoun *it*, provided this particular combination can be associated to a specific resolution path for the anaphoric reference. This category was included in the annotation because once there is a predictable resolution path to the collocations that appear in the corpus, chances for identifying the antecedent with a good degree of precision are extremely high.

The most typical collocation which appeared in the whole corpus was the cleft sentence (that-clauses). An example is given below as (10).

> (10) It [NA;Ex;Coll] is the author's experience *that children with these anterior open bites typically swallow with a tongue thrust.*

The collocation is the underlined part of the sentence above. A detailed discussion of collocation and the collocations for the English anaphors found in the occurrences is presented in Chapter 5.

3.5 – Standards for Correct Translation

This section describes the criteria used to analyze the translations performed by *Systran* and *Power Translator Pro*. They are in number of three and formed the main framework for translation analysis. In other words, all the occurrences were classified according to three categories, namely: *correct, incorrect* and *partially correct*.

The first and second ones were used to determine when an MT result was to be considered good or bad translation. Just for the present work and analysis I define as a bad (incorrect) translation, when the machine translator fails to translate the English pronoun *it* appropriately into Portuguese having the human-made translation as a parameter of comparison. The sentence structure was not taken into consideration, only the pronoun itself. The third category, *partially correct*, refers to cases in which the translation was not totally correct because of the verb that followed the pronoun. For instance:

(11) When there is *recession* and the soft tissue margin is alveolar mucosa, it [NA;Ex;Coll] is incorrect *to call it [Diff;Ex;DK] "gingival recession"*.

(11a) Quando há caída e a margem de tecido macia é mucosa alveolar, está incorreto **para** chamar isto "caída de gingival". *Power Translator Pro (PTP)*

(11b) Quando há recesso e a margem de tecido flexível é mucosa alveolar, é incorreto chamá-la de "recesso gengival". *Human Translation (HT)*

In the case above, the translation could be considered correct but since the program translates word-by-word, the preposition "para (to)" made the text unintelligible. Only a person who knows the English language or a program instructed with such knowledge could notice that what has been taken for the preposition actually refers to the infinitive form of the verb *to call*.

In short, as it was previously mentioned in this Chapter the annotation scheme is extremely important when working with the parallel corpus in this thesis. The annotation deals with difficult cases and prevents the researcher from putting them aside while encoding each occurrence.

The next chapters present the results of testing with existing software and the limitations found while they performed the translation task. The results for each software tested were cross tabulated against each of the properties in the annotation, so as to ascertain the relevance of the information contained in it for anaphora resolution.

CHAPTER 4

TESTING WITH EXISTING SOFTWARE

AND DIAGNOSIS OF LIMITATIONS

Poetry is what gets lost in translation.
Robert Frost (1874-1963)

As pointed out in Chapter 3, the data analyzed consisted of 305 tokens of *it* extracted from a 45,866-word corpus. All the occurrences were translated into Portuguese by two different machine translation systems and the results were later evaluated according to three categories, namely: *correct*, *incorrect* and *partially correct*, having the human-made translation as a parameter for comparison.

In a first moment, there was an attempt to work with three MT systems: *Systran*, *Power Translator Pro* and *Tradunet*[12] but, since *Tradunet* was very difficult and confusing to operate, it was discarded.

In the subsequent section I discuss features of the tested software before presenting the results of the evaluation.

4.1 – Features of Tested Software

In order to increase understanding of the way these systems operate, some features of each are described below. The selected systems are among the most commonly available in the international market and are routinely used by translation companies all over the world.

POWER TRANSLATOR PRO: It is a PC-based MT system. It is a commercial application for word-processor-generated texts, email and Web pages. *PTP* is described

in the user's guide as a package for producing 'draft translations that you can edit into final form'. It provides a bi-directional translation for a language pair. That means, it translates English to and from other 5 languages: French, German, Italian, Portuguese and Spanish

SYSTRAN: It is a MT system which runs on many different platforms such as Windows 95,98, NT and Linux. *Systran*'s methodology is described on its web site as a 'sentence by sentence approach', focused on individual words and their dictionary data, then on the parse of the sentence unit, followed by the translation of the parsed sentence. Sager (1993) states that *Systran* has been used "for the production of un-edited output, for post-editing use, for restricted-language document input. This system has the largest number of language pairs, all official European Community Languages, with varying sizes of dictionaries and states of quality of performance."[13]

From the systems mentioned above, *Systran* is the oldest and the best-known first generation MT system. It was developed in California – USA – in the 60's by the US Department of Defense and is currently being marketed by a French company. Some of their main clients are the Commission of European Communities, large commercial multinational corporations, as well as Internet enterprises such as Altavista, Alis/Netscape and others. There is a limited sample of the program available for prospective clients on the Internet[14].

One relevant aspect of each software concerns what companies say about their products and the translations they perform. Since these translation programs were designed to translate texts in a fast way, *Power Translator Pro* manufacturers suggest that it is necessary to pre-edit the source text for the machine in order to ensure a more accurate translation. This pre-editing stage[15] ensures that the vocabulary and structures

in the text match the translating capabilities of the MT system. It should be pointed out that *Systran* manufacturers, unlike the other software, published a short note in their website emphasizing that the program tries to achieve the highest possible accuracy in translation. They also claim that no translation is perfect, therefore their program is not intended to replace human professionals in this field. Moreover, they go on to state that users should note that the quality of the source text affects significantly the translations. The company also offers a human translation service to customers.

> Human Translation – This service is provided by *Systran* in alliance with Berlitz GlobalNet. If you want your document or file to be translated by a professional translator, please complete the following. Your translation will be sent back to your e-mail address within 48 hours for a document of 2,000 words or less. Please note that each translated document carries a minimum fee of US$ 50.[16]

There were some attempts to get more specific information about each MT system as to how the machine deals with the question of ambiguity; what the strategies used to translate the pronoun *it* are and if the programs were developed within a corpus-based approach. Unfortunately, these attempts to establish contact were not successful. *Systran* manufacturers just did not answer the e-mail message. Only after a second attempt to reach some information, Lernout & Hauspie, developers of *Power Translator Pro*, were willing to say something about their product. Nonetheless, they provided a very limited amount of information concerning the software (the e-mail messages are shown in Appendix 3). This influenced the analysis since the programs performances were observed and I am not sure how they have been programmed to deal with certain difficult resolution situations.

4.2 – Results of Testing

After all translated occurrences have been evaluated according to the *correct*, *incorrect* and *partially correct* categories mentioned previously in the beginning of this chapter, I could conclude that *Systran* works better than *Power Translator Pro* software.

Graph 4.1- Systran General Results
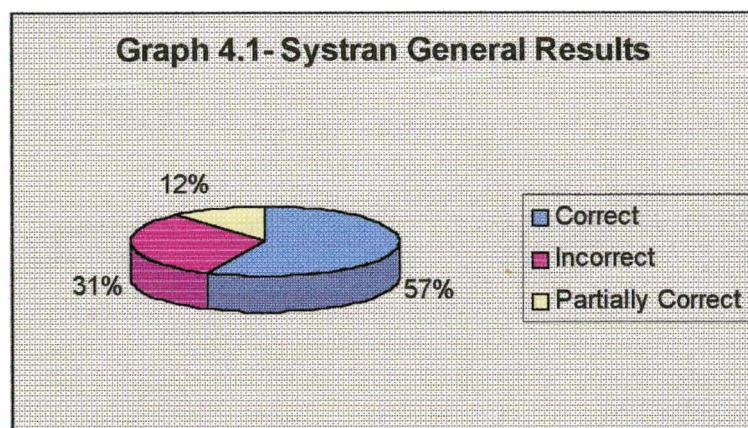
Systran presents 57% of correct cases, this means 174 out of 305 *it* tokens, while *Power Translator Pro* presents 51% (158 tokens).

Graph 4.2 - Power Translator Pro General Results
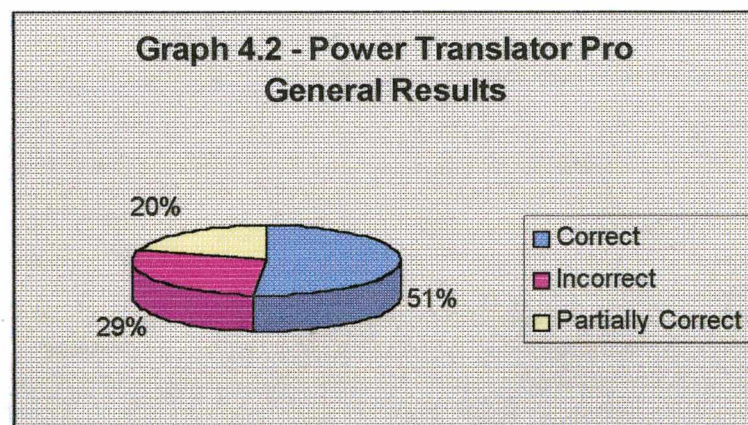
My conclusion does not rely only on these percentages since there is not a substantial difference between them. Both systems show almost the same results in Graphs 4.1 and 4.2. By this reason, the results were also all cross-tabulated, the *correct*, *incorrect* and *partially correct* categories for automatic translation results against each of the properties in the annotation scheme and later calculated with a chi-square test in

order to give a measure of how probable it was that a certain distribution of an independent variable across categories of a dependent variable was the result of chance.

All data concerning a given property is presented as a block. That means tables 4.1 and 4.2 refer to *Power Translator Pro* and *Systran* against compared syntactic function. Tables 4.3 and 4.4 concern *PTP* and *Systran* against type of antecedent. Finally, tables 4.5 and 4.6 show the processing strategy for the two software mentioned above against processing strategy.

It can be seen in table 4.1 that 70 (44.30%) out of the 158 correct occurrences are non-applicable (NA). As defined in Chapter 3, section 3.4.1.1, this category is used to classify cases in which the comparison of syntactic functions does not apply to the analysis of the anaphor. If this result is compared to table 4.2 we can perceive that both systems work best[17] in cases in which the syntactic function is non-applicable, when *it* is non-anaphoric. Almost all occurrences of NA tokens are resolved by means of collocation knowledge. Collocation, as it will be discussed in detail in the next Chapter in section 5.2, is being employed here/in this study as a perceptible combination of elements that comprises a token of the pronoun *it*, provided this particular combination can be associated to a specific resolution path for the anaphoric reference. For example:

> (12) In order to insure the viability of the cruise market going forward in addition to maintaining the loyalty of the existing cruise market. It [NA; Ex; Coll] is clear *that new cruisers will need to be drawn into the market*. To this end, it [NA; Ex; Coll] is important *to gain a better understanding of cruising prospects.*

If we contrast example (12) with the collocation list notion presented in Chapter 5, at least in what regards *it + X-be + Adj + that-clause* and *it + X-be + Adj + non-finite clause with to*, we can conclude that the collocation is the whole expression, including both anaphor and antecedent, in a relationship which is stable and predictable, at least in these two cases above.

Power Translator Pro also presents a high percentage of *partially correct* translations of *it* tokens in which the comparison of syntactic functions with the antecedent was non-applicable (NA) in the English text. When the results for *Different* and *Same* syntactic function referring to the *Incorrect* category were cross-tabulated, I noticed that *PTP* is not able to translate correctly cases in which the syntactic function of the pronoun *it* is different from its referent. Cases classified as NA are therefore easier to translate for *PTP*, very probably because of the omitting strategy which works invariably for the most common cleft sentences and other forms of sentential *it* (as defined in Quirk et al.).

Table 4.1 - Power Translator Pro Results by Compared Syntactic Function

	Same	Diff	NA	Total
Correct	56	32	70	158
Incorrect	44	36	7	87
Partially Correct	24	9	27	60
Total	124	77	104	305

Table 4.2 shows that the *Systran* made more errors in *Same* variable than in the other ones. A relative percentage of these errors when calculated represent 47.37% of the total incorrect cases. These cases are mostly related to discourse knowledge in which is necessary a fully processed discourse for the information required for the anaphor resolution.

Table 4.2 - Systran Results by Compared Syntactic Function

	Same	Diff	NA	Total
Correct	65	41	68	174
Incorrect	45	28	22	95
Partially Correct	14	8	14	36
Total	124	77	104	305

Through a comparison of tables 4.3 and 4.4, it is possible to see that there is not a great difference between the two programs performance. Both show almost the same results for *correct, incorrect* and *partially correct* categories. This mean that *Type of Antecedent* may be discarded in the annotation because is not significant to this research and does not affect in the anaphora resolution.

Table 4.3– Power Translator Pro Results by Type of Antecedent

	Ex	Im	NR	Total
Correct	140	6	12	158
Incorrect	81	4	2	87
Partially Correct	55	1	4	60
Total	276	11	18	305

Once antecedents are mainly *explicit*, this influenced considerably the results when comparing to the other two categories. Even though *implicit* cases seemed not so significant due the presence of just a few occurrences, my belief is that the results could change reasonably if they were analyzed in a larger corpus.

Table 4.4– Systran Results by Type of Antecedent

	Ex	Im	NR	Total
Correct	161	4	9	174
Incorrect	83	6	6	95
Partially Correct	32	1	3	36
Total	276	11	18	305

The third variable analyzed and presented in the annotation scheme organization is processing strategy. As it can be seen in tables 4.5 and 4.6 collocation knowledge seems to be relevant to translation since it is the category which both *PTP* and *Systran* present the best results. Collocations only seem to be difficult to translate in partially correct cases in one system.

The fact that the machine is able to translate collocations appropriately is not related to the knowledge of this property itself but to the pronoun omission, that is quite common for the systems, as it will be shown in the following pages. The next example illustrates the category with a translation of the token performed by *Systran*:

> (13) Pacifier-sucking is more common today than some fifteen-twenty years ago (Fig 1), but persists longer (Larsson, 1985; Larsson et al., 1992) and often occurs outdoors when the child is playing. It [NA;NR;Coll] seems as if *the pacifier* has been accepted as a part of the child's clothing, with which it [Same;Ex;DK] is often matched in color.

> (13a) Pacifier-sugar é hoje mais comum do que certos anos fifteen-twenty há (o fig 1), mas persiste mais por muito tempo (Larsson, 1985; Larsson et o al, 1992) e ocorrem frequentemente ao ar livre quando a criança está jogando. Parece como se o mais pacifier foi aceitado como uma peça da roupa da criança, com que é combinada frequentemente na cor. *Systran*

Analyzing the two tables for processing strategy, we can conclude that resolutions which require discourse knowledge processing are the most difficult ones. In *PTP* there are 42 tokens that represent 48.28 % of the incorrect anaphoric references while in *Systran* there are 40 tokens out of 107 which represent 42.11%. This results are due to the fact that processing strategies based on discourse knowledge demand more complex information which cannot be obtained in a collocation list or simply checking the semantic content of the anaphor. It requires the use of previously established discourse information for resolution, which is not present in the programs as can be noticed in the translated examples. Once each specific incorrect DK token is identified through a list of cases and strategies to resolve it are developed within the MT software, chances for successful automatic translation are high.

Table 4.5 - Power Translator Pro Results by Processing Strategy

	DK	FtC	FtCCh	Coll	Total
Correct	46	35	13	64	158
Incorrect	42	31	7	7	87
Partially Correct	19	14	3	24	60
Total	107	80	23	95	305

Comparing results for partially correct translated tokens, it can be seen that the totals are quite different from one system to the other. *Systran* seems to work better than *PTP*. *Systran* presents a relatively small percentage of cases that involve discourse and collocation knowledge and no first candidate chain cases.

Table 4.6 - Systran Results by Processing Strategy

	DK	FtC	FtCCh	Coll	Total
Correct	50	45	15	64	174
Incorrect	40	27	8	20	95
Partially Correct	17	8	0	11	36
Total	107	80	23	95	305

The next section presents the main difficulties found in the translated anaphoric cases as well as an attempt to organize better the numbers from section 4.2 so as to allow me to obtain more conclusions about the performance of the tested MT systems.

4.3 – Discussion

During the comparative evaluation of translation results between systems and human performance, some particular cases seemed to be relevant because they can explain why the systems are not efficient enough. As it was pointed out in the previous section, the MT systems are able to translate collocations correctly because of the omission of the pronoun *it* and not due to any kind of processing strategy. Unlike English, Portuguese is a language that often accepts this pronoun omission. So, if a system or human translator omits the pronoun there will be a great probability of

correctness in translation. On the other hand, there are cases in which the simple omission is not possible and the machine is not able to identify these cases, causing an incorrect translation. The machine is not capable of identifying these cases because the software is only programmed to work on specific situations and this may generalize the translation task. This means that whenever a pronoun appears in the subject position in both main and subordinate clauses, the MT system will automatically omit the pronoun. This omission is perfectly acceptable in most cases, yet it can sometimes result in inappropriate translation (see example 14).

> (14) Gargiulo, Wentz and Orban described *this dimension*. It [Diff;Ex;FtC] varies in width because of variation in width of the supra-alveolar connective tissue fibers.

> (14a) Gargiulo, Wentz e Orban descreveram esta dimensão. Varia na largura por causa da variação na largura das fibras supra-alveolar do tecido conexivo. *Systran*

> (14b) Gargiulo, Wentz e Orban descreveram esta dimensão. Varia em largura por causa de variação em largura das fibras de tecido conjuntivo supra-alveolares. *PTP*

In the example above, both systems omitted the anaphor. However, the best translation would be one in which the omission did not occur. In the example 14c, which is human translation, the antecedent is clearly and easily identified due to the use of the personal pronoun "ela".

> (14c) Gargiulo, Wentz e Orban descreveram esta dimensão. Ela varia em largura por causa da variação em largura das fibras do tecido conectivo supra-alveolar.HT

Another significant aspect when assessing the quality/correctness of automatic translation concerns lexical knowledge. The software presents difficulties in dealing with different meanings of a word and applying the one which best suits the context. Example 15 below shows a typical case in which this lexical difficulty is presented:

(15) It [NA; Ex; Coll] has been claimed *that supracrestal injuries to the dentogingival attachment reversible provided the lesions are allowed to heal against biologically acceptable root surfaces*.

(15a) Foi reivindicado que danos de supracrestal para o anexo de dentogingival reversível contanto as lesões são permitidas curar contra superfícies de raiz biologicamente aceitáveis. *PTP*

(15b) Reivindicou-se que os ferimentos supracrestal ao reversible dentogingival do acessório forneceram os lesions estão reservados para heal de encontro à raiz biológica aceitável aplainam. *Systran*

The inadequate use of a word can completely alter the meaning of a sentence or cause difficulties in its processing. The proper translation would be one in which the meaning is similar to that expressed by the sentence suggested below.

(15c) Afirma-se que danos supracrestais à ligação dento-gengival são reversíveis, desde que se deixe as lesões cicatrizarem em superfícies de raiz biologicamente aceitáveis. HT

Word choice is not the only lexical problem found during the evaluation. In certain cases already mentioned previously both systems failed to translate some words, which affected the rest of the paragraph. This difficulty is due to the fact that the programs are fed with a limited, non-technical dictionary.

The foregoing analysis of the performance of each program allows us to conclude that the machines do not seem to have any specific criteria to identify the antecedent in a sentence. This is reflected mainly in gender problems. Nevertheless, it is worth pointing out that this does not constitute a problem only for MT systems but for humans. There are discourse knowledge cases which are hard to deal with and which affect the translation task. These cases require the use of previously established discourse information for appropriate results. The following example can clearly illustrate this:

(16) When there is *recession* and the soft tissue margin is alveolar mucosa, it [NA;Ex;Coll] is incorrect *to call it [Diff;Ex;DK] "gingival recession"*. At the time of examination, one cannot be

certain whether the tissue margin was alveolar mucosa when the tooth erupted or whether it [NA;NR;DK] had been gingiva (free) which had been lost through recession and is presently alveolar mucosa. "Marginal tissue recession" is a more accurate description of the process that has occurred.

(16a) Quando há recesso e a margem de tecido flexível é mucosa alveolar, é incorreto chamá-la de "recesso gengival". No momento do exame, não se consegue saber com segurança se a margem do tecido era mucosa alveolar quando do nascimento do dente ou se tinha sido gengiva (livre) que tinha sido perdida por recessão e é, presentemente, mucosa alveolar. "Recesso de tecido marginal" é uma descrição mais exata do processo ocorrido. TH

Based on the annotation in the original example, it is clearly visible that the human translator was misled in the antecedent choice because of the ambiguity in the original text. The example 16a above serves to highlight the fact that even if the translation is performed by a human being this is not enough to assure a hundred percent accuracy.

Mitkov (1996) states that the establishment of antecedents of anaphors is indispensable for correct translation. Moreover, he also says that:

> When translating into languages which mark gender pronouns, for example, it is essential to resolve the anaphoric relation. On the other hand, anaphor resolution is vital when translating discourse rather than isolated sentences since anaphoric references to preceding discourse entities have to be identified. Unfortunately, the majority of MT systems do not deal with anaphor resolution and their successful operation usually does not go beyond the sentence level. (p.7)

Taking up the results discussed in the previous section and the points raised here, I can say that *Systran* translates anaphora better than *Power Translator Pro*. The former translated correctly 174 tokens out of 305, which represents 57.05% of the total. *Systran* also seems to use the Portuguese object[18] personal pronoun more often and more adequately than *PTP*, as the example below shows.

> (17) One day the Buddha was sought out by a follower who carried with him presents to show his devotion. The man stepped forward to the Buddha, held out his hand and offered *a priceless ornament.* The

Buddha said, "drop it [Same; Ex; DK]", The man, surprised, did and stepped back. Then he came forward again this time with *a precious jewel* in hand. The Buddha said, "drop it [Diff; Ex; DK] ". The follower did and stepped back. Then the man smiling stepped forward and held out both empty hands. The Buddha again said, "drop it [Diff; Ex; DK] ".

(17a) Um dia o Buddha foi procurado para fora por um seguidor que carregasse com ele os presentes para mostrar sua devoção. O homem pisou para a frente ao Buddha, prendeu para fora de sua mão e ofereceu um ornament priceless. O Buddha dito, "deixa-o cair", o homem, surpreendido, fê-lo e pisou-o para trás. Então veio para a frente outra vez esta vez com uma jóia preciosa à disposição. O Buddha disse, "deixe-o cair". O seguidor fez e pisou para trás. Então sorrir do homem pisado para a frente e prendido para fora de ambas as mãos vazias. O Buddha disse outra vez, "deixe-o cair". *Systran*

Testings using a statistical program showed up no great significance in the analysis of cross-tabulated data. The Chi-square test and the association measures[19] were not useful due to some cells having less then the minimum number expected for the program to operate successfully.

By observing the tables in the previous section, we can conclude that even though *implict* cases appeared in a very low amount compared to *explicit* ones, they seem to be really relevant in anaphora resolution. This assumption is based on the presupposition that *implicit* occurrences have a difficult resolution pattern. The program made 6 errors out of 11 total *implicits* (54.54%) for *Systran*. A careful study of this category in a large corpus is expected in future.

The error rate in cases that have been classified as resolved by the strategy named discourse knowledge in the present research is considerably high for both MT programs (*PTP* 48.28%, *Systran* 42.22%) and a more complex study of each individual case is needed. By means of a minute/conscientious analysis of each discourse knowledge incorrect occurrence and improvement of the problematic cases in terms of anaphoric resolution, the error rate can be reduced in order to achieve a more successful automatic translation.

CHAPTER 5

COMMENTS ON THE ANNOTATION SCHEME

IN THE ENGLISH CORPUS

> It is a capital mistake to theorize before one has data. Sir Arthur Conan Doyle (1859-1930)

As pointed out in Chapter 4, each case of anaphora was analyzed and annotated according to a set of three properties as an attempt to classify anaphora cases and identify patterns of anaphora resolution on the basis of this classification. Thus, assuming the automatic annotation of a text using the proposed scheme, it would be possible to identify antecedents and consequently improve machine translation results in what concerns the pronoun *it*, as initially hypothesized.

The first property which appeared in the annotation scheme order is *syntactic function*, in which the categories *same*, *different* and *non-applicable* are used to classify, in a traditional way, the syntactic function of each element in the anaphora relationship (that is, anaphor and antecedent), according to traditional syntactic functions within the English sentence. The second property marked in the annotation scheme is the *type of antecedent*, which is defined as *implicit*, *explicit* or *non-referential*. The type of antecedent is the identification of the type of referent within the text. The last property, which is called *processing strategy*, is an attempt to define the sort of knowledge used to resolve anaphora cases, incorporating psycholinguistic elements into the annotation.

5.1 – Statistical Analysis

It seems appropriate to initiate the discussion with the data pertaining to the syntactic function variable, since it is the first item analyzed in the annotation scheme

order. Graph 5.1 below shows the distribution of compared syntactic functions according to the classification previously defined in approximate percentages. It can be seen that, in 41% percent of the anaphoric references, which represent 124 tokens out of 305, the anaphor and the antecedent have the same syntactic function. In all cases analyzed, there was only one occurrence with a collocation structure[20] [Same;Ex;Coll], because this is not usual. On the other hand, the 34% percent representing the non-applicable function is mainly collocation since it is a characteristic of such a structure. A different function, which is 25% percent of the total (78 tokens), is divided in discourse knowledge and first candidate strategy .

Graph 5.1 - Compared Syntatic Function
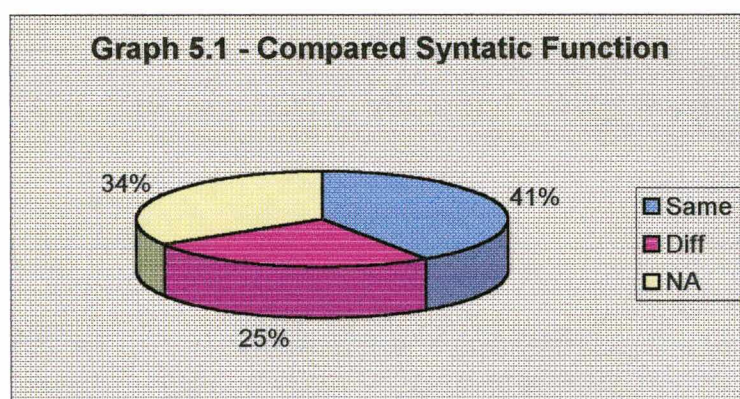
The second variable analyzed was the type of antecedent. Almost all tokens, which represent 276 out of 305, are explicit. The most interesting results are those related to discourse knowledge and collocations because from all these explicit cases, 31.41% are discourse knowledge and 28.16% are Collocations . Non-Referential cases represent 6% of the total, mainly with the collocation structure [NA;NR;Coll]. Almost all occurrences (94.44%) of non-referential pronouns are resolved by means of collocation knowledge. This means that a list of collocations which contain anaphoric words, together with basic resolution procedures based on the investigation of corpus material, provided by the computational linguist to the system designer could be extremely useful in anaphora resolution, especially if the collocations which contain

non-referential pronouns could be identified with a reasonable degree of precision. In contrast to that, Machine Translation systems actual performance are not good since they are not developed within a corpus based approach and do not have an accurate resolution pattern.

Numbers for implicit antecedents are even lower. There are only 11 occurrences out of 305 occurrences, which represents 4% of the total. As it was pointed out in the previous chapter, even though explicit occurrences take place in a great number and influenced statistical results, implicit cases seem to be worth studying due to their difficult resolution pattern.

Graph 5.2 - Type of Antecedent

4% 6%

90%

Ex
Im
NR
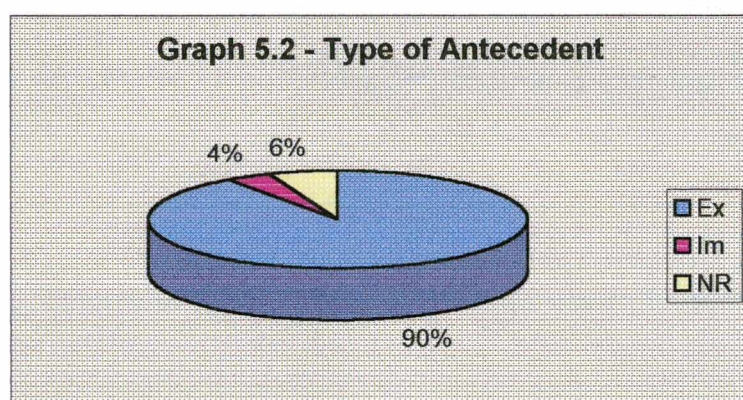
The third variable analyzed in the annotation scheme is the processing strategy. It is worth pointing out that the behavior of the pronoun *it* has inspired the creation of this variable. Graph 5.3 shows the distribution of the occurrences by the relevant categories with the approximate percentages. It is suggested that 35% percent of the anaphoric references have discourse knowledge as the most probable processing strategy. As mentioned in the previous chapter, this happens because discourse knowledge embodies a large amount of cases in which a full processing of discourse is necessary to identify the antecedent. This occurs when an intervening acceptable antecedent has to be ignored in favor of the intended one. Once each specific type of discourse knowledge is identified and stored in the MT program together with its

resolution path, chances of successful automatic translation are quite high. Let us consider the following example:

(18) Connective tissue apparently forms a capsule around *blade vent implants*, with fibers circumferentially forming around it [Same; Ex; DK] without attachment.

In the example above the antecedent is a plural noun phrase and not a singular noun as it is usually expected because of number agreement. The strategies to resolve this case are: to ignore agreement conflict, check lexical clues and use encyclopedic knowledge. Otherwise, only someone knowledgeable in dentistry is able to identify precisely the correct antecedent because for an ordinary/lay reader, it is not totally clear if *a capsule* or *blade vent implants* is the most appropriate antecedent.

The next example below contains two other different cases of discourse knowledge with both explicit and implicit antecedent:

(19) *Hearing*

This is almost certainly of tremendous importance for communication between members of a herd at all levels, on land or in water. On land it [NA;Ex;DK] is also used by bulls for territorial control, and by mother seals which probably identify the individual calls of their young before confirming it [NA;Im;DK] with smell.

The first *it* token above is referred to by Rocha (1998) as *candidate bypass*, the characteristic of candidate bypass is, as the name suggests, that a candidate for correct antecedent identified by means of a first-candidate strategy is to be bypassed. The checking of syntactic and lexical clues is a standard procedure in the identification of every antecedent. It should result here in the discarding of the first candidate. In the example 19 the antecedent is a title of a book section. In the second case, not only does the identification rely on lexical clues but it needs a more sophisticated processing strategy. This type of occurrence is called *strained anaphora* and it consists of a noun extracted from a verb. In the example above, the inferable antecedent which is derived

from identify is identification. It worth pointing out, that this type of discourse knowledge is strongly related to implicit cases.

Another type of discourse knowledge present in the corpus is the following one:

(20) Another method is to draw *a net* at sea and then drive the animals off the shore into it [Diff;Ex;DK].

In this example, bypassing the first candidate or lexical clues is not enough for the anaphora resolution. The human reader and the machine would probably think that *at sea* is the antecedent and not *a net*. One useful aid in such a situation is to use a syntactical analysis together with the items mentioned before.

Knowledge of collocations also scores high among the processing strategies for anaphor resolution (31%). Therefore, as it was mentioned previously, it seems interesting to provide potential/candidate anaphora interpreters with a list of collocations containing anaphoric words together with basic resolution procedures based on the data investigation in this thesis.

There are 80 occurrences out of 305 in which the first possible candidate available is the antecedent. Numbers decrease significantly when it refers to referential chains (23 tokens). Referential chain is a not typical type of processing strategy found in the corpus.

Graph 5.3 - Processing Strategy
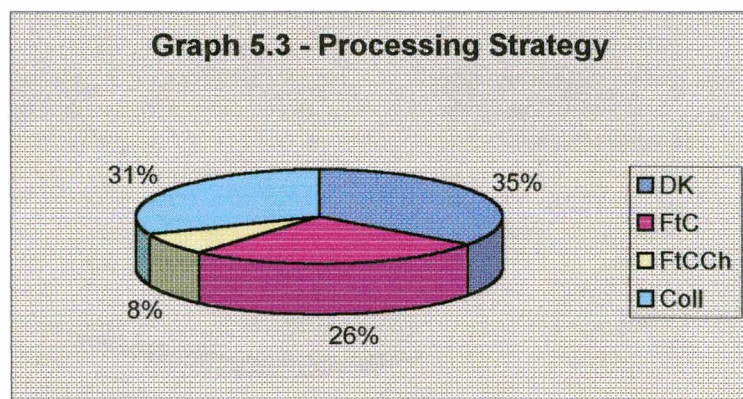
Graphics 5.4, 5.5 and 5.6 show the most regular structures and their possible ways of being grouped according the criteria elected by the analyst as well as their frequency in the whole corpus.

Graph 5.4 below shows that from all *Same* cases, the two most typical ones are those related to discourse knowledge and first candidate search. As it was mentioned in Chapter 3, first candidate search is quite common because the antecedent can easily be found by simply looking back in a sentence and taking the first and most probable candidate (see the following example).

(21) Restorative procedures must be delayed until a new gingival crevice develops after periodontal surgery. *This waiting period* may be 6 weeks, but it [Same;Ex;FtC] is usually much longer.

Graph 5.4 - Same
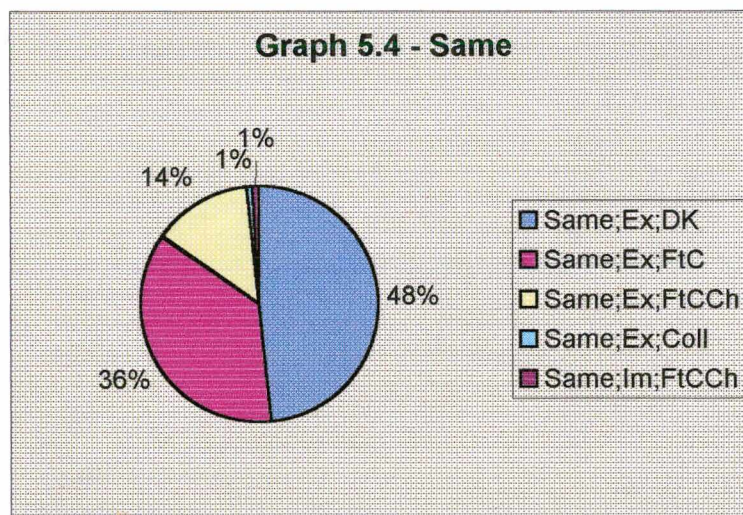
In Graph 5.5 almost all non-applicable structures are resolved by means of collocation knowledge. Anaphora cases classified as resolved by means of this particular strategy represent 90% of cases in which the comparison of antecedent-anaphor syntactic functions on the dichotomous same-different basis is non-applicable. Examples 22 and 23 below show anaphora cases of this kind in two observed combinations with the classification for type of antecedent. The first one contains a cleft sentence, in which antecedents - the subsequent that-clause - are invariably explicit. The

second one is a case of the so-called 'prop it' (Quirk et al.,1985), in which the *it* token is classified as non-referential, as it clearly has no antecedent and might be left out of the sample. For reasons previously stated, however (see 3.4.1.2), non-referential cases of the 'prop it' were invariably classified as such. No tokens of implicit antecedent are associated to the processing strategy based on collocation knowledge.

(22) It [NA;Ex;Coll] is common for trainers *to start with young animals, mainly for economic reasons*.

(23) It [NA;NR;Coll] was after 6 o'clock when I slipped back into the room, just as the early morning light was coming through the "Mrs. P," I reached out and touched her arm. She raised her tear streaked face to look at me. "It's [NA;NR;Coll] time," I said and waited. When she was ready, I helped her off the bed and held her in my arms for a few moments. We cried together. "Thank you, nurse," she said as she looked into my eyes and pressed my hand between hers. Then she turned and walked away.

A more elaborated discussion about collocations and a list of the most typical cases found in the corpus are presented in the next section.

Graph 5.5 - NA
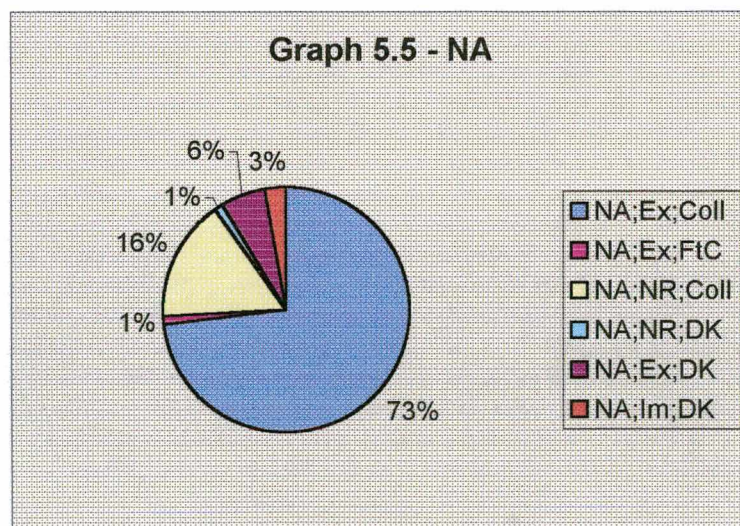
From all the syntactic function tokens analyzed in this research, *different* is the one which presents the lowest frequency.

Graph 5.6 - Diff
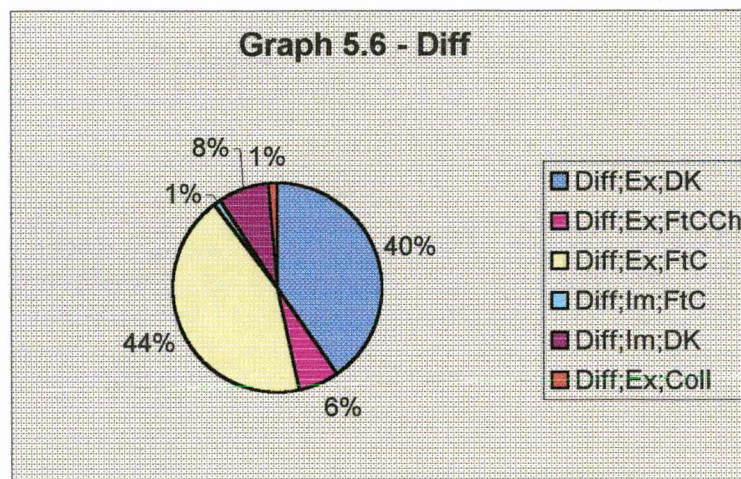
Table 5.1 summarizes the foregoing discussion by presenting a general cross-tabulation of the processing strategy against the results for type of antecedent and syntactic function in all the occurrences of *it*, in an attempt to find out more about the ways in which this particular anaphor is resolved. All cases have already been discussed at the beginning of this section.

Table 5.1 - General Cross-Tabulation of Data

Processing Strategy	Function									Total
	Same			Diff			NA			
	Ex	Im	NR	Ex	Im	NR	Ex	Im	NR	
DK	60			31	6		6	3	1	107
FtC	45			33	1		1			80
FtCCh	17	1		5			0			23
Coll	1			1			76		17	95
Total	124			77			104			305

5.2 – More about Collocations

For the present purposes, the concept of collocation adopted in this thesis as it was mentioned in Chapter 3, is defined as an identifiable combination of elements that contains a token of the pronoun *it*, provided this particular combination could be

associated to a specific resolution path for the anaphoric reference. The most typical

collocation pattern which appeared in the whole corpus was the cleft sentence. A cleft

sentence occurs when the speaker or writer wants to emphasize/highlight one segment

in a sentence, so it is divided into a separate clause. According to Quirk et al (1972):

> A special construction, which gives both thematic and focal
> prominence to a particular element of the clause is the cleft
> sentence, so called because it divides a single clause into two
> separate sections, each with its own verb. Most cleft sentence
> statements begin with the empty pronoun *it* followed by the verb *be*,
> which in turn is followed by the element on which the focus falls.
> (p. 951)

Let us consider the following example extracted from my corpus::

(24) It [NA;Ex;Coll] is conceivable *that a permanent insult to the gingival attachment may cause self-limiting damage and that, with time, a new level of bone, apical to a newly located supracrestal connective tissue attachment will be formed.*

The sentence with *it* gives special importance to the fact that it is acceptable that

the situation mentioned in the example above may occur. Variants of the *that-clause*,

using *who* and *which*, can also be found in the corpus.

5.2.1 – Some Specifications About the Collocation List for the English Sample

The collocation list that appears in this section is just referring to entries of the

pronoun *it* created on the basis of data. It should be emphasized that this list may

change in spoken language. This list attempts to generalize patterns of resolution found

to be regular in specific contexts. The entries are grouped in the list according to the

kind of anaphor involved. For instance:

it + X-be + (modif) NP (modif) + that-clause

- Cataphoric *it*

- Antecedent = that-clause

This entry is meant to identify occurrences of cleft sentences in which *it* is systematically resolved as a cataphoric reference to a that-clause. The symbol X means any inflected form of the verb, also including negative, interrogative and affirmative forms. Symbols appearing in subsequent entries also include noun phrase (NP), adjective (Adj), nonfinite (NF), prepositional phrase (PP), adverb phrase (AdvP), object phrase (ObjP), and verb phrase (VP).

Here is an example of the entry mentioned in the beginning of this section:

(25) It [NA;Ex;Coll] is the author's experience *that children with these anterior open bites typically swallow with a tongue thrust.*

5.2.2 - The Collocation List for the English Sample

Each collocation is followed by an example extracted from the corpus. The *it* tokens that are discussed in each case appear in red and the antecedent in italics and underlined.

Cleft Sentences

- The most common collocation

- Recognition

    - it + X-be + time and/or place expression + that-clause

    - it + (AdvP) + X-be + (modif) NP (modif) + that-clause

    - it + X-be + (modif) ADJ (modif) + that-clause

    - it + X-be + (AdvP) + (PP) + that-clause

    - variants of the that-clause possible

    * who- clause; which-clause; *that* omitted;

- resolution path

    - the that-clause is the antecedent

(26) It [NA;Ex;Coll] has been claimed *that supracrestal injuries to the dentogingival attachment reversible provided the lesions are allowed to heal against biologically acceptable root surfaces.*

it + X-be + (AdvP) Adj + NF-clause

- cataphoric

- antecedent = NF-clause

(27) The scene is simply a term used to describe all of the objects and text that will end up being a part of your creation. If these objects move over time, your scene will become animated. It's [NA;Ex;Coll] important *to remember that there are an infinite amount of directions you can look into the 3D world of Swift,* but there is only one view that really matters, the Viewport.

it + X-be + Adj (modif) (FOR ObjP) + NF-clause

- resolution path

  - the antecedent is the NF-clause

(28) In general pinnipeds live in groups of various sizes. It [NA;Ex;Coll] is common for them *to congregate on land for breeding.*

it + X-become + comp Adj + NF-clause

- cataphoric

- resolution path

  - the antecedent is the NF-clause

(29) Business process standards. Use of standards is an important strategic policy, as many companies today limit the number of intercompany formats they support. Given the success in developing and adopting global standards in specific narrow areas (e.g., EDIFACT), one might argue that it [NA; Ex; Coll] is becoming more difficult *to make "standards mistakes" than was the case several years ago.*

it + X-be + time expression

- time of the clock; day; month; year;

- after/before phrases

- nonfinite clause complement

- resolution path

  - non-referential

(30) It [NA; NR; Coll] was after 6 o'clock when I slipped back into the room, just as the early morning light was coming through the "Mrs. P," I reached out and touched her arm. She raised her tear streaked face to look at me. "It's [NA; NR; Coll] time," I said and waited When she was ready, I helped her off the, bed and held her in my arms for a few moments. We cried together. "Thank you, nurse," she said as she looked into my eyes and pressed my hand between hers. Then she turned and walked away.

it + X-seem + as if –clause

- variants such as: as though- clause, possible

- antecedent = as if-clause

  (31) It is sometimes suggested *that continental drift has also been a factor in the origins of migration.* However, it is certainly not *easy to explain migration in terms of contemporary function in all cases,* and it [NA;Ex;Coll] seems <u>as if it may have its origins in history.</u>

it + X-seem + (modif) Adj + that-clause

- cataphoric

- antecedent = that-clause

  (32 A more recent suggestion proposes that the stones are swallowed for stabilization and/or ballast. This theory currently commands much popular support but in view of the anatomical modifications made by evolution in allowing mammals to reconquer an aquatic environment, including structural changes such as the repositioning of lungs, it [NA;Ex;Coll] seems unlikely *that stabilization or ballast would rely on each individual swallowing sufficient stones instinctive behaviour does not command such finesse.* It may be argued that the animal can meter the quantity of stone in some way, and thus learn how many to swallow, but it seems improbable.

but it + X-Verb + adj

- anaphoric

- resolution path

  - begin with that-clause (which is the antecedent)

  (33)A more recent suggestion proposes that the stones are swallowed for stabilization and/ or ballast. This theory currently commands much popular support but in view of the anatomical modifications made by evolution in allowing mammals to reconquer an aquatic environment, including structural changes such as the repositioning of lungs, it seems unlikely that stabilization or ballast would rely on each individual swallowing sufficient stones instinctive behaviour does not command such finesse. It may be argued *that the animal can meter the quantity of stone in some way, and thus learn how many to swallow*, but it [NA;Ex;Coll] seems improbable.

it + (AdvP) + X-seem + NP + that-clause

- cataphoric

- the that-clause is the antecedent

  (34 It [NA;Ex;Coll] always seems a pity *that these aquatic creatures are usually exhibited by trainers on land where they are to say the least inhibited, instead of being allowed to perform in water where they could acquit themselves with honours*. Scientists also train sea lions to assist them in research work.

it + might + X-be + VP

- that-clause is the antecedent

- variants of modal verbs possible (e.g. should)

  (35) The need for orthodontic treatment in the 16 year-old former pacifier-suckers was less compared to children with no previous pacifier-sucking habit, and particularly compared to previous digit-suckers (Table1). Indeed, the arches of previous pacifier-suckers were well developed and crowding was rare. It [NA; Ex; Coll] might be speculated *that the muscle activity caused by pacifier-sucking during early childhood could be of some beneficial influence on the development of the arches*, as well as promoting the slight anterior rotation of the maxilla and the mandible.

that makes it + (modif) Adj + NF-clause

- the antecedent is cataphoric

  - variants of the that-clause possible

* which-clause;

(36) Automating tile interface with customers to facilitate purchasing and with suppliers to make them an important part of Du Pont's business system has given the company a competitive advantage; it has done so by raising switching costs and supporting structural changes that make it [NA; Ex; Coll] possible *to share information with existing businesses, introduce and control new distribution channels, and exert influence over suppliers.*

X-make + it + Adj + (AdvP)

- the antecedent is the Adverb Phrase

(37) Each keyframe has a Start Animation and a Stop Animation control. Within one Frame the controls are set to a value of one so there is no pause. Adjusting these controls allows you to stop one aspect of tin animation for as long as you want. For example, you may want a spinning text object to pause mid-spin so it's easier to read. We've used the universal 'Green equals Go. Red equals Stop" color scheme to make it [NA;Ex;Coll] obvious *when an object is going to move, rotate or change size,* and when it is going to stop doing those things.

X-be + it + okay + If – clause

- the if-clause is the antecedent

(38) When I returned to her room, Tracy was drifting in and out of sleep and her mother, still sitting on the, edge of the was fighting to stay awake. "Tracy", I called softly, is it [NA;Ex;Coll] okay *if we lie down on the bed with you?'* She opened her eyes and nodded I shifted Tracy's thin body over to the make more room for her mother, then lay down beside on the opposite side of the bed

it + X-mean + NF-clause

- the antecedent is cataphoric

(39) Certain obligations result from this belief -such as a commitment to know self and other as caring. We believe that each

person, throughout life, grows in their capacity to express caring. Caring is process and as such is continually unfolding. It is lived moment to moment. Our experiences present opportunities to grow in an understanding of what it [NA;Ex;Coll] means *to be human and live caring.*

X-find + it+ Adj + NF – clause

- sense of *X-think it is*

- the antecedent is cataphoric

(40) Cooperation of country managers is often key to successful execution of global initiatives. Without a shared global vision of the business, the CIO will find it [NA;Ex;Coll] impossible *to sustain a global IT strategy.* Even in the global enterprise, the central problem is that managers are generally responsible for their local authorities and not for those of other regions or functions.

CHAPTER 6

FINAL REMARKS

> This is it!
> David A. Johnston (volcanologist). (His last
> words when on duty on May 18th, 1980 at the
> time Mount St. Helen's volcano blew its top.)

This study was mainly concerned with the analysis of anaphoric phenomena in real-life written language and how MT systems deal with the problem of translating the pronoun *it* into Portuguese. As it could be noticed throughout this work, a number of aspects require a lot more research than it was possible to accomplish within the limits of the study. Especially regarding how the existing failures can be overcome to make the MT system more efficient.

The corpus used, which was collected as a sufficient sample of the written English language for this linguistic investigation, can probably be of great usefulness in future studies as well as the development of the annotation scheme that was based on Rocha's model for spoken language. It might be argued that some of his categories did not apply to written language or to this research. One of the reasons for such was because he investigated anaphoric phenomena as a whole. However, the present analysis was restricted to a specific anaphor.

Furthermore, the annotation scheme resulted in a statistical analysis that supported the creation of a collocation list[21] and obtainment of information derived from observation of corpus data. Although some people may argue that collocation is not so relevant for an exhaustive discussion in this thesis. The information provided by the list was relevant to emphasize the need of a corpus-based approach in MT. Another

important aspect found in the analysis was that even *implicit* cases seemed not so significant due the presence of just a few cases. Consequently, further research with a larger corpus is needed in this area of study. The error rate for cases classified as being resolved by means of the processing strategy named discourse knowledge was also significant for both Machine Translation programs since there was a high percentage of occurrences for them (*PTP* 48.28%, *Systran* 42.22%). A more complex study of each individual case is needed due to its relevance to anaphora resolution.

Nowadays, the new machine translation system that is emerging more powerfully than those mentioned in this thesis, for commercial purposes is a result of the influence of EBMT, for example: *TRADOS*[22]. This system is capable of storing through a 'translation memory' database a previously translated text and its source text. The stored information can be fast accessed and inserted in another translation work, saving the translator time and effort in translating more than once the same thing. This translated material can be stored in three levels: at sentence level, sentence part level and at terminology level.

One desirable proposal for scholars is to apply computational means in daily research activities. The use of computers in linguistics is so important as the use of a microscope, as a working tool, to the biology area. According to Leech (1992):

> the computer's ability to search, retrieve, sort, and calculate the contents of vast corpora of text, and to do all these things at an immense speed, gives us the ability to comprehend, and to account for, the contents of such corpora in a way which was not dreamed of in the pre-computational era of corpus.(p.106)

In conclusion, machine translation, although not yet mature, could be a tool of great value for translators. According to Crystal (1998) machines can undoubtedly help to take a great deal of the drudgery out of routine translation work. The use of on-line dictionaries, word processors, machine translation systems or translation aids can

enhance the productivity of translators. MT systems make human labor easier and provide a service that human beings cannot provide in terms of accuracy, consistency or precision. One reason why MT systems have not been accepted in actual translation environments is a severe shortage of formally trained specialists to perform the task of operating computational tools quite well.

Computational tools should be an active part in education, even in Junior or Senior High Schools. In fact, some places have computer hardware and software but people do not know how to use them appropriately. Having the equipment is one thing, using it well is quite another. If it is not used well, there is no educational value and it is just a poor employment of the computer resource. University students in Brazil (similarly to students in other countries such as the USA, at Carnegie Mellon University[23]) should be trained, while in the translation academic course, to use and research about such computational tools as MT because, through their feedback, developers of MT systems could have a guide to improve the existing technology.

NOTES

Chapter 1

[1] For more information about collocations, please refer to Chapters 3 and 5.

Chapter 2

[2] Chomsky is a professor of linguistics and currently involved in politics. In the 50's and 60's he revolutionized the work in linguistics with his theory of *Language Structure* known as *Transformational-generative grammar*. In his theory he works with two important notions/concepts: *surface structure and deep structure*. Crystal (1985) in his *Dictionary of linguistics and phonetics* briefly discusses these two terms. "The surface structure of a sentence is the final stage in the syntactic representation of a sentence, which provides the input to the phonological component of the grammar, and which thus most closely corresponds to the structure of the sentence we articulate and hear"(p. 297). The surface structure analysis does not explain cases of ambiguous sentences or how a person unthinkingly is able to relate sentences with different surface forms but with the same central meaning. As a consequence, the concept of deep structure was developed. It is "a level of structural organization in which all the factors determining structural interpretation are defined and interrelated. The standard view is that a grammar operates by generating a set of abstract deep structures, subsequently converting these underlying representations into surface structures by applying a set of transformational rules." (ibid)

[3] A thorough discussion of anaphoric relations will be presented in section 2.5.

[4] The term psycholinguistics is defined in the *Dictionary of Linguistics and Phonetics* as "a branch of LINGUISTICS which studies the correlation between linguistic behaviour and the psychological processes thought to underline that behaviour" (p.251).

[5] The *Information Technology Encyclopedia* (website), presents the following definition for Artificial Intelligence (AI): "the simulation of human intelligence processes by machines, especially computer systems. These processes include learning (the acquisition of information and rules using the information), reasoning (using the rules to reach approximate or definite conclusions), and self-correction. Particular applications of AI include expert systems, speech recognition, and image recognition."

[6] Lexical Functional Grammar (LFG): "A linguistic theory which claims that the SYNTACTIC STRUCTURE of a SENTENCE consists of a CONSTITUENT structure and a FUNCTIONAL structure, which represents superficial GRAMMATICAL RELATIONS. In this approach, the LEXICAL COMPONENT is assigned much of the role formerly associated with the syntactic component of a TRANSFORMATIONAL grammar" (Crystal, 1985, p.178).

Chapter 3

[7] For further details, please refer to section 3.5

[8] For details on Chi-square test, see Carlson and Thorne, 1997.

[9] Cataphora is the opposite of anaphora. Crystal (1985) defines the former as a "term used for the process or result of a linguistic unit referring forward to another unit. *Cataphoric reference* is one way of marking the identity between what is being expressed and what is about to be expressed" (p.43).

[10] The main candidate is determined through the processing strategy mentioned in section 3.4.1.3.

[11] For further details see Chapter 5 – section 5.2.

Chapter 4

[12] TRADUNET: It is an English - Portuguese translator available for almost all text editors and electronic mails as well as Microsoft Word for Windows (95, 98 or NT), version 7 or 97 (Portuguese or English) and WEB navigators (Netscape 4.0, Microsoft Internet Explorer, version 3.0 or 4.0 (Portuguese or English)). Developed in Brazil.

[13] Sager, J.C. (1993). The Pragmatic Circumstances of Automation in Translation. In J.C Sager, Language Engineering and Translation - Consequences of Automation.

[14] http://www.Systransoft.com

[15] There is a guideline to pre-edit the source text in the help file. The guideline is the following one: 1. Avoid idiomatic expressions, trying to express the same idea in a different way, 2. Use short sentences expressing the main idea instead of long sentences, 3. When writing dates, try to avoid using only numbers. The best option would be like this: October, 21-1999, 4. Write the text as clear as possible using a formal language.

[16] http://www.Systransoft.com

[17] It is relevant to say that the parameter used to determine when a program translates better a category than another consists of a comparison of percentages, line vs. column.

[18] According to Faraco & Moura (1998), the object personal pronoun works as an object or complement.

[19] Even *Goodman and Kruskal tau* which is a reliable association measure test based on the notion of proportional reduction of error, was not totally trustworthy in its results.

Chapter 5

[20] *Structure* in this thesis is a three-property combination such as [NA;Ex;Coll].

Chapter 6

[21] This collocation list when more developed containing anaphoric words, together with basic resolution procedures may be a useful tool to the anaphora interpreter.

[22] TRADOS translator's workbench was developed in Stuttgart, Germany. It is a database in which source - language sentences are stored together with their target-language equivalents. It avoids time-consuming and repetitive tasks. *Trados* manufacturers affirm that the program can achieve its task because "it stores all linguistics data in what referred to as neural networks. This powerful new technology enables it to look for all kind of data in a flash and with a degree of flexibility and assurance that convince you as you try the different functions of Translator's Workbench". (website)

[23] This University was the pioneer in creating a center for machine translation in cooperation with the school of computer science in the U.S.

REFERENCES

Carlson, W. L. & Thorne, B. (1997). Applied Statistical Methods. Prentice-Hall. USA.

Carreiras, M.; Oakhill, J. & Garnham, A. (1993). The Use of Superficial and Meaning-Based Representations in Interpreting Pronouns: Evidence from Spanish. European Journal of Cognitive Psychology, volume 5 (1), (pp.93-116).

Crystal, D. (1985). A Dictionary of Linguistics and Phonetics. (2nd ed.) – (The Language Library). Basil Blackwell Ltd. Oxford.

Crystal, D.(1994). Dictionary of Language and Languages. London, Penguin.

Crystal, D. (1998). The Cambridge Encyclopedia of Language. C.U.P.

Dyvik, H. (1995). Exploiting Structural Similarities in Machine Translation. Computers and the Humanities, volume 28, (pp.225 – 234).

Francis, W.N (1992). Language Corpora B.C. In I. Svartvik (ed.), Directions in Corpus Linguistics. Proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991. (pp.17-31) Monton de Gruyter: Berlin.

Gale, W.A. & Church, K.W. (1993). A Program for Aligning Sentences in Bilingual Corpora. Association for Computational Linguistics, volume 19, number 1, (pp.75-90).

Harris, T.L. & Hodges, R. E. (eds.) (1995). The Literacy Dictionary: The Vocabulary of Reading and Writing. Newark, International Reading Association.

Hoey, M. (1991). Patterns of Lexis in Text. Describing English Language. Series Editors: J. Sinclair & R. Carter. Oxford University Press.

Information Technology Encyclopedia [website] http://whatis.techtarget.com.

Jones, D. (1996). Analogical Natural Language Processing. Studies in Computational Linguistics. Series Editor Harold Somers.

Leech, G. (1992). Corpora and Theories of Linguistic Performance. In I. Svartvik (ed.), Directions in Corpus Linguistics. Proceedings of Nobel Symposium 82, Stockholm, 4-8 August 1991. (pp.105-122) Monton de Gruyter: Berlin.

Lewis, D. (1996). The Power Translator: an Evaluation of a PC-based MT System. Machine Translation Review, Issue 4, (pp.22 –32).

Liceras, J. M. (1989). One Some properties of the "prodrop" parameter: looking for missing subjects in non-native Spanish. In S. Gass & J. Schachter, Linguistic Perspectives on Second Language Acquisition. (pp. 109 – 116) CUP – Cambridge.

Mitkov, R. (1996). Anaphora and Machine Translation. Machine Translation Review, Issue 4, 6 –16.

Newton, J. (1992). The Perkins Experience. In John Newton (ed.) Computers in Translation: A Practical Appraisal, London:Routledge. (pp. 46 - 57).

Quirk,R.; Greenbaum,S., Leech, G. & Svartvik, J. (1972). A Grammar of Contemporary English. Harlow, Longman.

Quirk, R.; Greenbaum, S., Leech, G. & Svartvik, J. (1985). A Comprehensive Grammar of the English Language. London: Longman.

Richards, J.; Platt, J.& Platt, H. (1992). Longman Dictionary of Language Teaching & Applied Linguistics. Harlow, Longman.

Rocha, M.A.E. (1997). Supporting Anaphor Resolution in Dialogues with a Corpus-Based Probabilistic Model. In R. Mitkov & B. Boguraev (eds), Operational Factors in Practical, Robust Anaphora Resolution for Unrestricted Texts. Proceedings of a workshop sponsored by the Association for Computational Linguistics. Madrid, Spain.

Rocha, M.A.E. (1998). Corpus-Based Study of Anaphora in Dialogues in English and Portuguese. Tese de doutorado, Brighton: University of Sussex,

Sager, J.C. (1993). Language Engineering and Translation - Consequences of Automation. John Benjamins Company - Amsterdam/ Philadelphia.

Sampson, G. (1987). MT: A nonconformist's view of the state of the art. In M. King (ed.), Machine Translation Today: the state of the art (pp.91-108). Edinburgh: Edinburgh University Press.

Shuttleworth, M. & Cowie, M. (1997). Dictionary of Translation Studies. Manchester, UK/St. Jerome Publishing.

Sinclair, J. (ed.) (1990). Collins Cobuild English Grammar. London, Harper Collins.

Sinclair, J. (1991). Corpus, Concordance, Collocation. Describing English Language. Series Editors: J. Sinclair & R. Carter. Oxford University Press.

Somers, H. (1997). A Practical Approach to Using Machine Translation Software: post-editing the source text. The Translator, volume 3, number 2, (pp.193-212.)

Swan, M. (1981). Practical English Usage. Oxford, Oxford University Press.

Systran [translation software]. http://www.systransoft.com.

Van Eynde, F. (ed.)(1993).Linguistic Issues in Machine Translation. Communication in Artificial Intelligence Series. Printer Publishers.

Waltermann, D. (1993). Machine Translation Systems in a Translation Curriculum. In C. Dollerup & A. Lindegaard (eds), Teaching Translation and Interpreting 2 - Papers from the 2nd Language International Conference. (pp. 309-317) John Benjamins Company - Amsterdam/ Philadelphia.

Whitelock, P. & Kilby, K. (1995). Linguistic and Computational Techniques in Machine Translation System Design. 2nd edition. Studies in Computational Linguistics. Series Editor Harold Somers.

APPENDIX 1

CORPUS ANNOTATED.

Artificial Sucking Habits: Etiology, Prevalence and Effect on Occlusion

Erik Larsson, O.D., Dr.
University of Lund, Sweden and Orthodontic Clinic, Falkoping, Sweden

It [NA;Ex;Coll] has been suggested (Larsson, 1985) *that there are three different types of sucking habits: pacifier-sucking, digit-sucking and unrestricted breast-feeding or non-nutritional suckling.* Unrestricted breast-feeding or suckling means sucking on the mother's breast for comfort, physiological contact, and satisfaction of the sucking urge, but not primarily to get milk.

The !Kung San infants are not able to cling so the mother uses *a sling*, which gives her maximal non-restrictiveness, leaving her arms and legs free. It [Diff;Ex;DK] allows skin-to skin contact between mother and infant and keeps *their infant* on the mother's hip, allowing it [Same;Ex;DK] to see what the mother sees. The infant has constant access to the mother's breasts, which are uncovered, and feeds whenever he or she likes.

It [NA;Ex; Coll] should be kept in mind *that the pacifier is in the child's mouth before the eruption of the teeth. The pacifier*, as well as the digit, does not create the anterior open bite, normally associated with the habit. It [Same;Ex;DK] does hinder the full eruption of the incisors and the vertical growth of the alveolar process (Larsson and Ronnerman, 1981).

It [NA;Ex;Coll] is the author's experience *that children with these anterior open bites typically swallow with a tongue thrust.*

Pacifier-sucking is more common today than some fifteen-twenty years ago (Fig 1), but persists longer (Larsson, 1985; Larsson et al., 1992) and often occurs outdoors when the child is playing. It [NA;NR;Coll] seems as if *the pacifier* has been accepted as a part of the child's clothing, with which it [Same;Ex;DK] is often matched in color.

The pacifier is not often sucked, but simply stays in the mouth passively. When playing, for instance, it [Same;Ex;FtC] is a good way for the parents to prevent the child from putting objects into the mouth. Most children are so used to *the pacifier* that they are capable of speaking intelligibly with it [Same;Ex;FtC] in the mouth (Lindsten et al.)

The need for orthodontic treatment in the 16 year-old former pacifier-suckers was less compared to children with no previous pacifiersucking habit, and particularly compared to previous digit-suckers (Table 1). Indeed, the arches of previous pacifier-suckers were well developed and crowding was rare. It [NA;Ex;Coll] might be speculated *that the muscle activity caused by pacifier-sucking during early childhood could be of some beneficial influence on the development of the arches*, as well as promoting the slight anterior rotation of the maxilla and the mandible.

The main reason, from the orthodontist's point of view, for accepting *pacifier-sucking* is that it [Diff;Ex;FtC] reduces the number of children, who develop a digit-sucking habit (Larsson and Dahlin, 1985; Larsson, 1985).

Those dealing with pacifier-sucking, should attempt to prevent the child from using the pacifier incorrectly. Many of the problems described above can be controlled by educating the parents today many children have the pacifier in their mouth almost full-time. In these cases, it [NA;Ex;Coll] should be pointed out *that pacifier-sucking is a substitute for non-nutritive feeding or unrestricted breast feeding and not a physiological part of the baby's growth.*

In those rare cases, in which it [NA;Ex;Coll] is decided *to try to break a pacifier-sucking habit,* the following method is recommended. The therapist, the child and the parents decide to break the habit on a special day some months ahead. A suitable day to choose is the last school day before a holiday. The child selects 4 - 5 small presents for the parents to buy.

In the anteroposterior direction, the effect of digit-sucking in a young child is not especially noticeable. Often there is only a slight spacing in the incisor region and a slightly increased overjet. This effect increases with age and, In the permanent detention, the effect of continued digit-sucking is usually quite obvious.
It [NA;Ex;Coll] is useful *to distinguish between two types of digit-sucking: Typical thumbsucking or sucking with the ventral side of the fingers facing the palate and the maxillary incisors; and sucking with the dorsal side of the fingers upwards.* In the latter case the finger or fingers lie passive and the effect should be similar to that of pacifier-sucking.

There are, of course, psychological reasons for stopping a sucking habit. A 10 year-old digitsucker is normally embarrassed because of *the habit* and is anxious to hide it [Diff;Ex;FtC] from friends.

Restorative Dentistry and the Periodontium

Robert I. Sachs, D. D. S.

A patient who has *this problem* will usually be quite familiar with the specific site or sites where it [Diff;Ex;DK] occurs.

Food impaction is defined as "the forceful wedging of food through occlusal pressure into the interproximal space." It [Same;Ex;FtC] can occur because of open proximal carious lesions open proximal contact areas rotations and malpositioning of teeth wear of teeth and improper restoration of tile top of the tooth (Fig. 1).

The dentist is sometimes faced with a situation in which a restoration needs to be placed oil the proximal surface of a tooth adjacent to a narrow space. Should the restoration be overcontoured to attempt closure of the space? Loose contacts seem to be more related to food impaction than wide open contacts, so it [NA;Ex;Coll] is probably better *to leave a space wide open than to flail to establish a tight contact all other things being equal.*

If the floss pulls through the contact area with decreased resistance as the patient clenches, the occlusal force is wedging the teeth apart. It [NA;Im;DK]is then a matter of identifying the contacting inclines and reshaping by selective grinding to eliminate the wedging force.

Overcontoured full crown restorations are usually the consequence of a chain of events beginning with underpreparation of the tooth, proceeding to the laboratory technician deciding that an esthetic color from an adequate veneer thickness is more important than proper axial contours and ending with the clinician accepting *the overcontoured crown* and placing it [Same;Ex;FtC] on the patient's tooth.

When the margin is located on enamel the sulcus will be wider than when the margin is located on the anatomic root it [NA;Ex;Coll] is important *to preserve the original sulcus morphology in the intrasulcular contours of the restoration.*

Invasion of the subsulcular space is best prevented by avoiding the placement of intracrevicular margins. When esthetics dictate such a need collarless crowns should be considered. Meticulous attention to tooth preparation and gentle impression procedures will prevent inadvertent encroachment, but when it [NA;Ex;Coll] is determined *that a restoration must enter the subsulcular space,* a crown lengthening procedure is indicated (Fig. 16).

Branemark coined the word osseointegration to describe the intimate relationship that lie observed occurring between bone and metal of his screw-shaped titanium implants." The self-tapping design provides initial stabilization of *the implant,* which is submerged below the gingiva for 3 to 6 months while bone remodels around it [Diff;Ex;DK].

Connective tissue apparently forms a capsule around *blade vent implants,* with fibers circumferentially forming around it [Same;Ex; DK] without attachment.

Marginal Fit of Gold Inlay Castings

GORDON J. CHRISTENSEN, D.D.S., M.S.D.
University of Washington, School of Dentistry Seattle, Wash.

FOR MANY YEARS the gold inlay has been regarded by dentists as one of the finest restorations for treating moderate to large carious lesions. However, it [NA;Ex;Coll] has also been accepted *that a marginal opening in gold inlays is unavoidable because of the film thickness of the cementing medium.*

The teeth were mounted so that they could be taken out by removing a section of orthodontic wire which ran through a hole in the root of *each tooth* to hold it [Diff;Ex;DK] in position.

An investing and casting technique which insured undistorted, bubble-free, well-fitted castings was necessary. When *an unpolished casting* was tried on a tooth, and it [Same;Ex;DK] showed an obvious marginal discrepancy, the casting was remade because attempts to burnish or "spin" gold into voids in excess of the slightest visually detectable space usually end in mutilation of the tooth and an unavoidable cement line in the final restoration.

Those who disfavor the gold inlay should evaluate qualitatively the marginal adaptation of their finished gold inlays, and if it [NA;Ex;Coll] is found *that this quality could be improved,* a change in finishing technique is suggested.

Periodontal Response to Long-Term Abuse of the Gingival Attachment by Supracrestal Amalgam Restorations

Tal H, Soldinger X Dreiangel A and Pitaru S: Periodontal response to long-term abuse of the gingival attachment by supracrestal amalgam restorations. J Clin Periodontal 1989; 16: 654-659.

After bone loss, root surfaces which were previously attached to alveolar bone by periodontal ligament were (0.90 mm) attached to connective tissue. Mild inflammatory infiltrate was observed in the lamina propria primarily at the restoration area. It [NA;Ex;Coll] is concluded *that continuous abuse of the gingival attachment results in a certain loss of the periodontal attachment apparatus.*

It [NA;Ex;Coll] has been claimed *that supracrestal injuries to the dentogingival attachment reversible provided the lesions are allowed to heal against biologically acceptable root surfaces* (Loe & Silness 1963).

It [NA;Ex;Coll] is conceivable *that a permanent insult to the gingival attachment may cause self-limiting damage and that, with time, a new level of bone, apical to a newly located supracrestal connective tissue attachment will be formed.*

Based on the present results, the minimal length of the "biologic width" which is compatible with the biologic and mechanical needs of the periodontium in the beagle, averages 0.90 mm (Table 1). This size is below the 4.47 min observed in the control specimens. It [NA;Ex;Coll] is not clear *whether 0.90 mm is the ultimate size of the supracrestal attachment or whether the process of bone resorption would have continued had the animals not been sacrificed at 1 year.* Based on previous data from short-term studies (Parma-Benefati et al. 1985, Tal & Stahl 1986, Tal et al. 1988), it [NA;Ex;Coll] is conceivable *that no further significant bone and attachment loss arc expected and that the periodontium may function with a reduced gingival attachment.*

Physiologic Dimensions of the Periodontium Significant to the Restorative Dentist

By J. GARY MAYNARD, JR., D.D.S.
RICHARD DANIEL K. 'WILSON, D.D.S.

These components have been termed the physiologic dimensions of the periodontium that are fundamental to successful restorative dentistry. *The word* "physiologic" was chosen because it [Same;Ex;FtC] means consistent with the normal functioning of an organism.

This classification serves as a guide and teaching aid to the practitioner. It [Same;Ex;FtC] encompasses previously recognized landmarks and structures of the periodontium and attempts to organize them in a, fashion that is meaningful to the restorative dentist.

SUPERFICIAL PHYSIOLOGIC DIMENSION

The choice of the word superficial was made because this dimension is obvious to the clinician when he examines the tooth and its surrounding tissue. *The superficial physiologic dimension* extends from the mucogingival junction to the gingival margin (Fig. 1). It [Same;Ex;DK] is composed of the attached gingiva and the free gingiva.

MARGINAL TISSUE RECESSION

When there is *recession* and the soft tissue margin is alveolar mucosa, it [NA;Ex;Coll] is incorrect *to call it [Diff;Ex;DK] "gingival recession"*. At the time of examination, one cannot be certain whether the tissue margin was alveolar mucosa when the tooth erupted or whether it [NA;NR;DK] had been gingiva (free) which had been lost through recession and is presently alveolar mucosa. "Marginal tissue recession" is a more accurate description of the process that has occurred.

CREVICULAR PHYSIOLOGIC DIMENSION

The gingival crevice extends from the free gingival margin to the junctional epithelium (Fig. 1). Normally it [Same;Ex;DK] has depth that measures from 0 to 3 to 4 mm and it [Same;Ex;FtCCh] is lined with thin crevicular epithelium.

Depth of the crevice can be described as being either excessive or adequate for restorative dentistry. Excessive crevicular depth is characteristic of periodontal disease. Diagnosis of periodontal disease is made by " . . . inspection, palpation, and probing. *The periodontal pocket* is pathognomonic of periodontal disease; it [Same;Ex;DK] is present in no other disease.

Restorative procedures must be delayed until a new gingival crevice develops after periodontal surgery. *This waiting period* may be 6 weeks, but it [Same;Ex;FtC] is usually much longer.

Gargiulo, Wentz and Orban described *this dimension*. It [Diff;Ex;FtC] varies in width because of variation in width of the supra-alveolar connective tissue fibers.

When treating patients, the objectives of restorative therapy must be clear. The first and most basic objective is preservation of the teeth. *The attainment of this objective* would be less complex if it [Same;Ex;DK] could be considered independent of restoration of function, comfort and esthetics, but such is not the case. The latter objectives usually require sophisticated restorative dentistry and often include restorations with intracrevicular margins. Although it [NA;Ex;Coll] is widely accepted *that the best restorative margin is one that is placed coronal to marginal tissue*, most restorations have margins in the gingival crevice, and permanent tissue damage is common.

CLIA Cruising: The Best Vacation
1996 CLIA Ratings by Travel Agents

In order to insure the viability of the cruise market going forward in addition to maintaining the loyalty of the existing cruise market. It [NA;Ex;Coll] is clear *that new cruisers will need to be drawn into the market*. To this end, it [NA;Ex;Coll] is important *to gain a better understanding of cruising prospects.*

• Most prior cruisers plan to cruise again within the next five years. *Cruising* is such a positive experience that it [Same;Ex;DK] generates interest in repeating the experience.

• *Interest in cruising* is highest among prior cruisers, but it [Same;Ex;FtC] is not restricted to this group. Many of those who take other types of vacations intend to take a cruise. Vacationers who spend money on other types of vacations will be converted to cruising more readily than non-vacationers or those who take non-paying vacations (i.e.. stay with friends or relatives or take a vacation as part of a business trip.)

CRUISE PROSPECTS

* *This research* identifies five distinct segments of cruise prospects who differ in terms of who they are, what they expect from their vacations, and what prevents them from taking cruises. These segments present different opportunities for the cruise market. It [Same;Ex;DK] also suggests different marketing messages to appeal to each segment and to convert the prospects into cruisers.
The five segments, which are explained later, are: Family Folks, Want-it-Alls, Adventurers, Comfortable Spenders and Cautious Travelers.

{CONSIDERED TAKING A CRUISE

While the large majority of cruise prospects (83%) have thought about *taking a cruise*. fewer than one out of four (22%) have seriously considered it [NA;Ex;DK].

61% N-39 million	=	Never thought about taking a cruise
(17% N= 11 million	=	Thought about *taking a cruise* but did not seriously consider it)
[Diff;Ex;DK]		
22% N= 14 million	=	Seriously considered taking a cruise}

The following segment profiles represent -tvpical- descriptions of segment members. They reflect the central tendencies of each segment and highlight how the segments differ from each other. It [NA;Ex;Coll] should be noted, however, *that variations do exist within each of the segments and there are some commonalities across segments.*

Messages
- Cruises offer the highest standards of quality
- Cruises offer it [Diff;Ex;DK] all: *luxurious accommodations. excellent food. nightlife*

Network Management Strategy and Architecture

A network architecture defines where applications are executed and databases are located and identifies the communications links needed among locations; it [Same;Ex;DK] also sets standards to ensure that other ISA components are interrelated and work is important for providing standards for interconnecting very different systems instead of enforcing communality among systems.

Business process standards. Use of standards is an important strategic policy, as litany companies today limit the number of intercompany formal they support. Given the success in developing and adopting global standards in specific narrow areas (e.g., EDIFACT), one might argue that it [NA;Ex;Coll] is becoming more difficult *to make "standards mistakes than was the case several years ago.*

One central policy governing key standards, (e.g., EDI, SOL) should include a management agenda for promoting understanding of both tile standards and the standards setting rig process within industry, national, and international environments. Such a *central policy* accomplishes several objectives: it [Same;Ex;FtC] reduces costs, avoids vendor viability, achieves economies of scale, reduces potential interface problems, and facilitates transborder data flows.

Many CIOs do not understand how IT can transform operations globally. Their entrenched, nonintegrated business processes and systems because they were not designed to be cross-functional, much less cross-national, naturally resist a global IT strategy. Moreover, old systems carry the baggage of diverse technology and incompatible applications, which makes it [NA;Ex;Coll] even more difficult *to build common global systems.* A major New York bank that *began to replace incompatible systems with a global network* in the mid-1970s is still at it [Diff;Im;DK].

Cooperation of country managers is often key to successful execution of global initiatives. Without a shared global vision of the business, the CIO will find it [NA;Ex;Coll] impossible *to sustain a global IT strategy.* Even in the global enterprise, the central problem is that managers are generally responsible for their local authorities and not for those of other regions or functions.

Du Pont's data management strategy. Du Pont plans to consolidate six data centers worldwide into four-two in the United States, one in Singapore, and one in Europe. By limiting its technology platforms, *the company* is able to remotely support its data centers and thereby achieve much greater flexibility in moving systems around the world. For example, it [Same;Ex;DK] operates its European data center (a clone of the major data center in the United States) without, systems programmers.

Automating tile interface with customers to facilitate purchasing and with suppliers to make them an important part of Du Pont's business system has given *the company* a competitive advantage; it [Diff;Ex;DK] has done so by raising switching costs and supporting structural changes that make it [NA;Ex;Coll] possible *to share information with existing businesses, introduce and control new distribution channels, and exert influence over suppliers.*

Norwegian customs department, for example, is significant in two respects. One, it

[Same;Ex;FtC] was a very early arrival oil the trade-related EDI scene, having come on line in August 1988, four months before the Singapore TradeNet system. And, two, in as much as it [Same;Ex;FtCCh] is not a comprehensive trade-related system, but rather one that focuses only oil customs, it [Same;Ex;FtCCh] represents only an intermediate solution to the trade documentation problem.

Such _a GIS management strategy_ is appropriate today because it [Same;Ex;FtC] facilitates coordination among a firm's value-chain activities and business units, and provides tile firm with tile flexibility and coordination necessary to deal effectively with changes in technologies and market structures. It [Same;Ex;FtCCh] also aligns information systems management strategy with corporate business strategy by providing a foundation for all information technology architecture.

Globalization, together with tile competitive posture of tile global enterprise, defines the competitive posture or nations and city-states. The issues related to coordination and control in of the alignment of its cross-industry competitive posture. It [NA;Ex;Coll] is incumbent on governments _to seek the level of intervention that strengthens the competitive position of the state in the global business community._

Nursing as Comforting: The Artistry of Caring

It [NA;Ex;Coll] is a privilege *to participate with you in this First Brazilian Conference on Caring Comforting a Nursing Perspective for the Third Millennium.* The Conference theme expresses a commitment to those we nurse. Historically, the role of the nurse as comforter is singled out by Florence Nightingale herself. "It [NA;NR;Coll] must never be lost sight of what observation is for. It [NA;NR;Coll] is not for the sake of piling up miscellaneous information or curious facts, but for the sake of saving life and increasing health and comfort" (p. 70)

Nursing textbooks consistently identify comfort as a goal of care. It [NA;NR;Coll] would seem that as nurses we would have a clear and common understanding for the meaning of the work comfort.

However, as I prepared this paper I became aware of the rich and varied meanings of comfort in nursing. *Comfort* has been used as a verb meaning causing or strengthening; as a noun it [Same;Ex;DK] is desired effect of being nursed; and as a process it [Same;Ex;FtCCh] refers to soothing or comforting.

Certain obligations result from this belief -such as a commitment to know self and other as caring. We believe that each person, throughout life, grows in their capacity to express caring. *Caring* is process and as such is continually unfolding. It [Same;Ex;DK] is lived moment to moment. Our experiences present opportunities to grow in an understanding of what it [NA;Ex;Coll] means *to be human and live caring.*

It [NA;Ex;Coll] is especially through our relationships *that we come to know self and other as caring and to draw forth caring possibilities.*

Another assumption of this theory is that personhood is the process of living grounded in caring. *Personhood* is living who we are. It [Same;Ex;FtC] means living life in congruence with one's beliefs and values. It [Same;Ex;FtCCh] means being authentic. Personhood is enhanced through participating in nurturing relationships with others

I'm an old woman now and nature is cruel
Tis her jest to make old age look like a fool
The body, it [NA;Ex;FtC] crumbles, grace and vigour depart

Because of the nurse's desire to know the other as caring person, the nurse intentionally enters the world of another and begins to understand calls for nursing. The nurse, through genuine openness and active patience, hears calls for nursing. These calls are for specific forms of caring that acknowledge, affirm, and sustain the other as they live caring uniquely. *The nurturing response* of the nurse to these calls is unique as it [Same;Ex;DK] reflects the beauty and wholeness of the individual nurse.

Living caring is a reciprocal process. It [Same;Ex;FtC] requires the personal investment of both caring persons - the nursed and the nurse who come to know each other as living caring through a mutual unfolding. Trust and courage are needed for such presence to occur. This mutual process is set in motion when the nurse risks entering the other's world and the other invites the nurse into 6 his;her sacred space. They share a coparticipative relationship in which each affirms and describes the caring that is

created in the moment. We refer to this encountering as the "caring between". Even though *this process* may appear so simple, we all know that it [Diff;Ex;FtC] can only occur in nursing practice situations which are intentional.

It [NA;Ex;Coll] is within the context of the nursing situation *that calls for comfort are heard.* The nurse through reflective practice gathers every cue, every hunch, every sign, and drawing on a vast knowledge base, reflects on the meaning of these in each particular situation.

Knowing person as whole is in opposition to much of what is seen today where the predominate model for being in relationship is analytical and reductionist. Through genuine dialogue, one becomes aware of the uniqueness of other. Because one accepts the other for who she;he is, it [NA;Ex;Coll] is possible *to truly direct what is said to the person that is and to affirm personhood.*

The nurse and nursed brine, to this lived experience the fullness of their being their wholeness. It [NA;Ex;Coll] is within the nursing situation *that the nurse fully engages in the process of coming to know other as caring person.* The nursed experiences the sincere presence of the nurse.

It [NA;Ex;Coll] is by being present to the wholeness of other *that calls for nursing are heard.* When deliberate presence occurs, *the Wholeness of being is felt.* It [NA;Ex;DK] involves a way of listening and communicating through which one gives of self When the nurse is with the nursed in a way which is nurturing and healing, he;she can support and sustain the nursed so they can let go of the pain. Being with the wholeness of other assists the other to let go. As Taylor (1992) states the "shared sense of humanity helps a nurse stand back quietly, to move in to help where it [NA;Ex;Coll] is appropriate, *to speak plainly and sensitively, to provide comfort at subtle levels and to be oneself in spite of professional knowledge and skills"* (p. 42). Through such presence the nurse expresses humility and active patience that allows the person the freedom to be one's self. Consciousness is raised through this experience adding to the moral basis for being in relationship. The story of the Buddha captures the essence of presence.

One day the Buddha was sought out by a follower who carried with him presents to show his devotion. The man stepped forward to the Buddha, held out his hand and offered *a priceless ornament.* The Buddha said, "drop it [Same;Ex;DK]", The man, surprised, did and stepped back. Then he came forward again this time with *a precious jewel* in hand. The Buddha said, "drop it [Diff;Ex;DK]". The follower did and stepped back. Then the man smiling stepped forward and held out both empty hands. The Buddha again said, "drop it [Diff/Ex/DK]".

Each person prepares to enter nursing situations in one's own way. For some, *meditation* may be the way of finding within self quiet solitude. It [Same;Ex;DK] may be a word or a phrase which has particular meaning to you and allows you to focus. According to Nouwen, it [NA;Ex;Coll] is in stillness *we listen. The solitude* of these moments fosters compassion and deepens an understanding of connection with other. It [Same;Ex;DK] prepares us to hear calls for caring and to live our caring expressions. "If we keep faith with that we truly hear, at our Center, in our ear's ear, we may better serve an ideal of quest and compassion" (Richard, p. 117).

There is no right or wrong to see in this _situation_ – it [Diff;Ex;FtC] is unique. As we listen, we might individually consider the calls for nursing we hear.

When I made rounds; Tracy was alone She looked mom pale mid fired than I'd ever seen her before As we chatted and I did my usual "nursey" things with vital signs and IV adjustments an unexplained feeling told her mother should be there. I even felt some resentment toward her mother creep in "Tracy?" I asked, "do you mind if _I call your mother tonight?_" She looked away, but said it [NA;Ex;DK] would be okay I Probably should ham picked lip some cue; but I didn't.

Tracy was awake Men her mother and I went into the room I wasn't prepared for the interaction between them I watched Mrs. P. stop about three feet horn the end of the bed and stand still Neither of them mid much while I was in the room. I left to answer a can bell and returned perhaps 15 minutes later to find the scene virtually unchanged. The distance and silence made me feel uncomfortable. I felt confused and so helpless it [NA;Im;DK] wasn't at all what I expected between a mother and her dying daughter.

"Mrs. P?" I asked, hesitantly, "won't you sit on the bed with Tracy and me?" "I don't think we're allowed", she said. _I closed the door and pulled the curtain around the bed_ as I replied, I think it [NA;Ex;DK] will be all right. Why don't you sit here and I'll sit on Tracy's other side" I steered her to the edge of the bed where she sat rather stiffly looking ill-at-ease. I sat on the other side and began stroking Tracy's arm, talking softly to her. Later, as Mr. P. began to relax, I was relieved to see her reach out to touch Tracy. I left them for a while to make rounds again on the other children. When I returned to her room, Tracy was drifting in and out of sleep and her mother, still sitting on the, edge of the was fighting to stay awake. "Tracy", I called softly, is it [NA;Ex;Coll] okay _if we lie down on the bed with you?'_ She opened her eyes and nodded I shifted Tracy's thin body over to the make more room for her mother, then lay down beside on the opposite side of the bed

I left them alone for a while to attend to my other mounting responsibilities. The medical item having been wakened by another call, inquired about her _vital sign_ "I don't know and it's [Diff;Ex;DK] not important right now," I remember cling him, "just leave her alone." I don't know how he knew riot to argue, but he did not press the issue and left.

It [NA;NR;Coll] was close to five in the morning when I returned to the room. Tears came to my eyes as I stood at the foot of the bed and saw Tracy wrapped in her mother's arms, their bodies pressed closely together. Mr. P. lifted her head up from the pillow when I approached the side of the bed to adjust the IV.

The routine of postmortem cam was all too familiar to - remove IV, identification tags on wrist and toe, body to the morgue within the hour it [Same;Ex;DK] all came clearly to my mind step by step, Her mother was still looking at me. "Please don't take her yet," she pleaded with me," please let me stay with her a while longer." I left the room and closed the door quietly behind me.

It [NA;NR;Coll] was after 6 o'clock when I slipped back into the room, just as the early morning light was coming through the "Mrs. P," I reached out and touched her arm. She raised her tear streaked face to look at me. "It's [NA;NR;Coll] time," I said and waited

When she was ready, I helped her off the, bed and held her in my arms for a few moments. We cried together. "Thank you, nurse," she said as she look into my eyes and pressed my hand between hers. Then she turned and walked away.

We are fortunate that Gayle shared this beautiful story with us. As *a story*, it [Diff;Ex;FtC] is now accessible to others for reflection on action. Stories evolve from reflection on practice. Such examples from nursing practice stimulate us to think about those nursing situations which are an important part of who we are as nurses.

Study of the artistic expressions visible through the medium of story brings to life the meaning of nurse as special artist and nursing as art. It [Same;Ex;DK] is "in the realm of the aesthetic, the nurse is free to know and express the beauty of the caring moment... it [NA;Ex;Coll] is this full engagement within the nursing situation *that the nurse truly knows caring in nursing*" (Boykin and Schoenhofer, p. 21).

"*Nursing* is an art; and if it [Same;Ex;FtC] is to be made an art, it [Same;Ex;FtCCh] requires as exclusive a devotion, as hard a preparation, as any painter's or sculptor's work; for what is the having to do with dead canvas or cold marble, compared with having to do with the living body - the temple of God's spirit? It [Same;Ex;FtCCh] is one of the Fine Arts; I had almost said, the finest of the fine Arts"

Pinnipedia: Seals, Sea Lions and Walruses

In general pinnipeds live in groups of various sizes. It [NA;Ex;Coll] is common for them *to congregate on land for breeding.*

The respiratory and cardiovascular systems, dealt with fully under separate headings, have been modified to enable *the animal* to dive for long periods. Since it [Diff;Ex;FtC] spends much of its life, and finds its food, in water, it [Same;Ex;FtCCh] is a fast and agile swimmer (see Locomotion).

BEHAVIOUR

Unlike the Cetacea, most pinnipeds come onto land to breed. At these times scientific observations on their social and sexual relationships can be made with relative ease *our knowledge* of their behaviour between haul out periods is as scanty as it [Same;Ex;DK] is of cetaceans in general.

Copulation occurs at varying times after parturition; in some species females come into oestrus soon after giving birth while in others it [Same;Ex;DK] is delayed until the pup is weaned. The interesting phenomenon of delayed implantation occurs in the Pinnipedia. *The fertilized egg* develops to the blastocyst stage when it [Same;Ex;DK] stops growing and is not implanted onto the wall of the womb.

Birth of young pinnipeds takes place with a minimum of assistance from the mother who seems almost indifferent to the procedure and makes no attempt to break *the umbilical cord.* It [Diff;Ex;FtC] usually severs as a result of the movements of both participants.

As soon as *the pup* is born it [Same;Ex;FtC] begins to seek-Out the nipple. The first attempts are usually vague and poorly directed but when it [Same;Im;FtCCh] finds its goal the milk is squirted into its mouth. One of the most interesting features of pinniped lactation is the long periods, Up to 2 weeks, which the pup can spend without nourishment of any kind while the mother is away hunting. In spite of this, pinniped pups gain weight at a remarkable rate due to the very high fat content of the milk. The stomachs of most pups contain stones and it [NA;Ex;Coll] has been argued *that these in some way reduce the hunger pains during the periods of starvation.* However, since most adults also ingest stones and sand, the explanation seems unlikely. The age of pups *when they take to the water* varies greatly from the birthday to several weeks afterwards depending on the species. It [NA;Ex;DK] seems to happen without the mother's involvement.

The physical limitations of pinniped pups make play a difficult undertaking on land. In spite of this, while their mothers are absent, they do indulge with their fellows in rather cumbersome activities. By nature inquisitive, they investigate their environment and mouth strange objects. It [NA;NR;Coll] is not, however, until they take to the water that the mammalian need for play can be fully seen.

A little is known of the migrating habits of the Pinnipedia. Some species migrate many thousands of miles each year, while others do no more than move around an island chain. The purpose of *migration* is only poorly understood and it [Diff;Ex;FtC] may

have evolved as a result of changes in environment.

It [NA;Ex;Coll] is sometimes suggested *that continental drift has also been a factor in the origins of migration.* However, it [NA;Ex;Coll] is certainly not *easy to explain migration in terms of contemporary function in all cases,* and it [NA;Ex;Coll] seems as if it [Diff;Ex;FtC] may have its origins in history.

The ability of pinnipeds to learn tricks is well known to circus enthusiasts although, until very recently, there has been little attempt to study this facility scientifically. One study found that they. compared favourably with cats and primates, and were considerably more intelligent than rodents. It [NA;Ex;Coll] is to be anticipated *that with the interest now being taken in marine animals by the United States Navy,* more information on the mental capacity of these creatures will be forthcoming.

CAPTIVITY
Another method is to draw *a net* at sea and then drive the animals off the shore into it [Diff;Ex;DK].

Captive pinnipeds are fed on a diet of fish; the California sea lion, Zalophus californianus, will consume between 4 and 14 kg of fish daily. It [NA;Ex;Coll] has been found *that some fish contain a substance which destroys thiamin one of the B vitamins.* Animals fed exclusively on fish containing this thiaminase may well show symptoms of thiamin deficiency. It [NA;Ex;Coll] is essential therefore *to offer a variety of fish,* including species which do not contain thiaminase, and to supplement the diet with thiamin. Mackerel, Scomber scombrus, for example are thiaminase free.

Of the phocids the harbour or *common seal,* Phoca vitulina, is most commonly kept although it [Same;Ex;FtC] is an animal which is not a favourite among zoo directors for exhibition.

CARDIOVASCULAR SYSTEM

The structure and function of pinniped hearts are very similar indeed to the ordinary mammalian pattern. *The arterial system* is likewise unmodified. It [Same;Ex;FtC] is in the venous system that most of the adaptations have occurred.

The venous system has considerable modifications which relate to the diving behaviour of the order. The jugular veins, so important in land mammals, have been much reduced. Most of the blood from the head passes into 2 hypocondylar veins which unite to form a sinus within the spinal canal. From this sinus, vessels pass dorsally on either side of the cervical vertebrae to form a very large vein the extradural vein. *This large single vessel* runs along above *the spinal cord* the length of the body as far as the sacrum where it [Same;Ex;DK] divides into two. Its importance to the vascular modification can be judged by its very large size and because of its frequent connections with all other parts of *the venous system.* It [same;Ex;DK] almost certainly acts as a rapid transporter of blood since it [Same;Ex;FtCCh] has no valvular structure.

Blood from the kidneys is drained into the stellate plexus which covers the kidneys. From here it [Same;Ex; DK] passes through one of 3 vessels to the branch of the posterior vena cava adjacent to it. [Same/Ex/DK]

This sphincter controls the flow of blood to the heart. In some species it [Same;Ex;DK] is a specific bundle or muscles, in others it [Same;Ex;FtCCh] is less well developed, being formed of strands of diaphragm muscle.

From the extradural vein, branches leave in the anterior abdomen to form the right and left azygos veins. In the chest they become the intercostal veins and the left azygos is reduced, its *blood* passing into the right azygos vein from where it [Same;Ex;DK] drains into the anterior vena cava. (See also Diving.)

DIGESTIVE SYSTEM

The stomach is a simple single curved chamber, it [Same;Ex;DK] is glandular and receive quantities of digestive juices and hydrochloric acid, which start the digestive process. *The duodenum* follows the stomach from which it [Same;Ex;FtC] is separated by the pyloric sphincter. The duodenum indistinctly joins the rest of the long small intestine. At the junction between the large and small intestines is a small, sometimes hardly visible caecum.

A more recent suggestion proposes that the stones are swallowed for stabilization and;or ballast. This theory currently commands much popular support but in view of the anatomical modifications made by evolution in allowing mammals to reconquer an aquatic environment, including structural changes such as the repositioning of lungs, it [NA;Ex;Coll] seems unlikely *that stabilization or ballast would rely on each individual swallowing sufficient stones instinctive behaviour does not command such finesse.* It [Na;Ex;Coll] may be argued *that the animal can meter the quantity of stone in some way, and thus learn how many to swallow,* but it seems improbable.

The presence of stones in the forestomach of the Cetacea, it [NA;Ex;Coll] has been suggested (see Digestive System, Cetacea) acts as a grinding mechanism for food reminiscent of the grit used by birds. Clearly if this holds true for Cetacea and birds, why not for pinnipeds? Experimental work with birds, however, shows that grit is not essential for digestion.

Young pups still being suckled by their mother are known to have stones in their stomach. Again, since they are often suckled at infrequent intervals, it [NA;Ex;Coll] has been suggested *that stones alleviate hunger pains.* A more acceptable explanation is that, as a result of normal exploratory behaviour, the pups mouth stones and accidentally swallow them. Clinical experience with dogs reveals that those individuals which play with stones not infrequently have several stones in their stomach. It [NA;Ex;Coll] seems most likely *that the habit of playing with stones learnt as a seal pup continues into adult life,* and with a system of deglutition so well designed for swallowing large objects a few stones inevitably find their way to the stomach.

DRINKING

The obvious question which arises in animals which spend their lives in and around *sea water* is, 'Do they drink it [Diff;Ex;FtC] If not, where do they get their fluid replacement?' The answer is that we do not know in the case of pinnipeds. They have not been seen drinking sea water but that would not be an easy observation to make.

They certainly will and do drink fresh water in preference to seawater in the unusual environmental conditions that exist during captive travel. On some experimental diets in captivity they also drink fresh water. In the wild however it [NA;Ex;Coll] may be *that they obtain sufficient moisture from their food.*

LOCOMOTION

Water with its great resistance presents problems for things which move even slowly through it [Diff;Ex;DK].

The various groups of pinnipeds adopt different methods of locomotion. The sea lions, fur seals and walrus all belong to the superfamily Otarioidea. Their fore limbs are long and paddle shaped and these are used as the main propulsive force in water. The hind limbs can be turned forwards and used to walk o or run on land. The surfaces of all flippers are either naked or at best sparsely covered with short hair. The neck is long and able to turn in water in order to search for food. Modifications have been made to the fore flipper to improve its propulsive efficiency. Since it [NA;Ex;Coll] is the surface area pushing through the water which provides locomotion, evolution has increased this by webbing the feet and extending the flipper beyond the bony digits. Strengthening fibrous bands also give the edge of the flipper rigidity. While swimming the hind flippers are positioned with soles together, forming a single functional rudder. The hind flippers are also provided with extra surface area by cartilaginous extensions to the digits and webbing. When moving on land the fore flippers are placed sole downwards on the ground, the longitudinal axis of the flipper at right angles to the body. When moving slowly the flippers are placed alternately. But when running they are brought forward together. The structure of the walrus is slightly different. *The flipper* is somewhat smaller in surface area and added power is provided by the lateral movements of the hind end of the body. In this it [Same;Ex;DK] is somewhere between the true seal on the one hand and the fur seal and sea lion on the other.

Phocid propulsive power comes from the hind part of the body, the flippers merely add manoeuvrability. Mechanically *a long neck,* as seen in the otariids, would hamper movement and it [Same;Ex;FtC] is therefore reduced in length. On land the phocids are, to say the least, at a physical disadvantage. Species differ in their ability to move out of water. Many adopt the 'inch worm' technique. They anchor the front end of the body at the sternum and draw the hindquarters forward; then, anchoring the pelvic region, they extend the body forward. Some species make use of their flippers, others do not. One interesting variation is found in *the crabeater seal,* Lobodon carcinophagus, which moves on land very effectively by using the same lateral movements of the hind part of the body as it [Diff;Ex;DK] does in water.

RESPIRATORY APPARATUS
The trachea divides into 2 bronchi which lead to the 2 lungs. In some species *the branching* occurs just before the lungs while in others it [Same;Ex;FtC] occurs closer to the larynx and the 2 long bronchi continue posteriorly, lying close together until they enter the lungs.

The lungs are placed dorsally in the chest to help buoyancy. *The diaphragm* is slightly modified in position, being more oblique than in land carnivores. It [Same;Ex;DK] has the normal attachment from the xiphisternum but curves backwards to the second

lumbar vertebra.

SENSES
Sight

The walrus pupil on land takes *the form* of a horizontal slit, and in the bearded seal it [Diff;Ex;DK] is diagonal. In spite of the importance of sight, blind animals exist in nature and it [NA;NR;Coll] seems that other senses can be developed to compensate.

Hearing

This is almost certainly of tremendous importance for communication between members of a herd at all levels, on land or in water. On land it [NA;Ex;DK] is also used by bulls for territorial control, and by mother seals which probably identify the individual calls of their young before confirming it [NA;Im;DK] with smell.

The ear structure is similar to that of other mammals. Since the meatus is closed when the animal is in water it [NA;Ex;Coll] is likely *that sounds are transmitted along the wall of the meatus as in the cetaceans.*

Smell

Smell, judging by the size of the olfactory lobes of the brain, is relatively unimportant. Mothers, however probably make positive identification of their offspring on the breeding grounds by smell. It [NA;Ex;Coll] also seems likely *that smell is used as a sexual signal during the breeding season.*

TRAINING
Although spasmodic attempts to train other species have been made with varying degrees of success, it [NA;NR;Coll] is the California seal lion, Zalophus californianus, *which for many years has been a favourite in circus rings.* In particular trainers exploit the long mobile neck which is ideally suited for balancing. No *performance* of trained sea lions is worth the name unless it [Same,Ex;DK] includes a group of these splendid creatures balancing large spinning balls on their noses.

It [NA;Ex;Coll] always seems a pity *that these aquatic creatures are usually exhibited by trainers on land where they are to say the least inhibited, instead of being allowed to perform in water where they could acquit themselves with honours.* Scientists also train sea lions to assist them in research work.

It [NA;Ex;Coll] is common for trainers *to start with young animals, mainly for economic reasons.*

It [NA;Ex;Coll] has always been the belief of experienced trainers *that young animals learn a trick faster if allowed to observe an older trained animal performing it* *[*Same;Ex;FtC*] first.* For many years the scientific world remained sceptical but recent scientific investigations have added weight to these empirical beliefs.
Although some taxonomists, particularly in the past, have placed pinnipeds as a suborder of the Carnivora, it [NA;Ex;Coll] is now usual *to consider them as a separate order in their own right.*

SWIFT3D   User guide

The scene is simply a term used to describe all of the objects and text that will end up being a part of your creation. If these objects move over time, your scene will become animated. It's [NA;Ex;Coll] important *to remember that there are an infinite amount of directions you can look into the 3D world of Swift,* but there is only one view that really matters, the Viewport.

The what?

The Viewport is your window into this 3D world you are about to create. Some like to Imagine it's [Same;Ex;Coll] like looking through a camera and whatever you see through the lens is what viewers will be seeing when you finally export your image We like this Imagery so we have named this the Camera View. You should be very interested in what at goes on within the confines of *the Viewport* since it's [Diff; Ex; Coll] where the only visible action will take place in the scene.

In our daily 3D world it's [NA;Ex;Coll] easy *to stay oriented since we have familiar reference points like walls,* floors and ceilings to figure out the respective locations of objects such as chairs, plants and people. Within Swift 3D's world, there are no objects until you create them, so we have supplied you with a 3D grid based on the X, Y and Z coordinate system.

Camera view

This little drop-down menu allows you to change your Camera View, or how you look into your scene. If you imagine a transparent cube surrounding your scene, you can view your creation from any one of the six sides of the cube. So your options are Front (default), Back, Top, Bottom, Left and Right. You may also notice that the right view is grayed out and not an option since *this view* is taken by the Secondary Camera when it [Same;Ex;DK] is shown. (More on that later in this chapter.)

Layout (Yes, I stuttered intentionally)

You can adjust the size of the Viewport by modifying the default dimensions. *This feature* is important since it [Same;Ex; FtC] determines the dimensions of your scene when you get to the export stage You can change these specifications at any point before exportation, but having a general idea of what sue and shape you want your finished product to be will help in the design process.

Camera

The Lens Length setting behaves just like a camera lens. The longer the lens length, the more the camera will zoom in on your scene. The shorter the lens length, the more distant you appear to be from your scene. There are other ways of *adjusting* how close your scene is to the viewer, but this setting allows you to do it [Diff;Ex;DK] without moving the camera. One thing to note is that, just like a camera lens, if you shorten the length, distortion occurs as your view becomes more panoramic in nature.

Outline displays objects as wire frames only, with no fill patterns. *This mode* can be

useful when you have complex 3D models and intricate animation paths since it [Same;Ex;DK] speeds up the process of redrawing your objects every time you make a change in your scene.

Box just shows the smallest three dimensional box that each of your objects could fit into. I'm not exactly sure why you'd use *this feature,* but the engineers said it [Diff;Ex;FtC] was useful so I didn't argue.

Environment

You may be asking yourself what the difference is between background and ambient light color. Or maybe you're not. I was a bit confused but a simple explanation cleared it [NA;NR;Coll] up.

Background

The Background setting controls what you see behind your objects, no matter what direction you look into your scene. It [Same;Ex;DK] has no effect on any of your objects and it does not get imported into Flash. However, you will see the background if you play your 3D SWF file In the Flash Player.

Cone - Places a hot fudge Sunday into your scene. (Oh, sorry. I was getting hungry. It's [NA;NR;Coll] actually just a cone.)

Torus - Places *a torus* into your scene. (I won't scoff if you call it [Same;Ex;DK] a doughnut, I do.)

Manipulating Objects

Selecting objects

To select *an object,* simply click on it [Diff;Ex;FtC] and look for a wire frame box to appear around it [Same;Ex;FtCCh].

Scaling button

Reset Transformation button

This feature allows you to undo changes you've made to an object. By selecting an object and clicking the Reset Transformation button you will move *the object* back to the 0.0,0 coordinates and remove any scaling and;or rotation you've applied, basically returning it [Same;Ex;DK] to its original location and size Any materials applied will remain and any animations that affect the object will also stay unchanged.

If you want to get technical, *the camera* is not panning around from a fixed point. It's [Same;Ex;DK] actually always pointing straight ahead and just being moved left, right, up and down

Camera Pan button

This button is only necessary when there is no background visible to select within the Viewport that does not contain an object, the main situation when this would occur is if an object is very close to the camera. With the Camera Pan button toggled on, the camera will pan no matter where you click and drag on the scene. It's [NA;Ex;Coll] impossible *to select objects in this mode.*

It's easy *to turn the Camera Pan button on and forget about it* [Diff;Ex;FtC]. Later, you're on the verge of a nervous breakdown because you can't select the flipping torus you just created There's enough things in the world to get frustrated about so I avoid this button at all cost.

Secondary Camera button

It's [NA;Ex;Coll] not always easy *to tell where your objects are in relation to one another along the Z axis,* making the Secondary Camera function very handy. It's [NA;Ex;Coll] also useful when you're creating animation paths for the same reason - *getting a better grip on depth.* When you adjust the view of *the secondary camera* by panning or zooming, it [Diff;Ex;FtC] does not affect the primary camera's view, and thus the scene. However, any movement or scaling of objects via the secondary camera Viewport will also take effect in the primary camera's view.

Crystal Trackballs

Crystal Trackballs are tools that allow you to position both objects and lights. They work like you were turning a virtual ball with your mouse. Just click and drag on the surface of *the ball* and it [Diff;Ex;FtC] will turn whichever way you choose. They're slick little units and quite easy to use. Let's take a look at each one individually.

Object Trackball

The Object Trackball is designed to let you adjust the orientation of objects in your scene. It [Same;Ex;DK] remains inactive until you have selected an object.

To move the lights around your scene you can select *a light* with a click and it [Diff;Ex;DK] will become a red wire frame. Now you can use the Lighting Trackball to position *the light* exactly where you want it [Same;Ex;FtC]. Use the Subtract Light button to remove a selected light.

When you want to rotate an object in a specific direction you can use these shortcuts to lock *the object trackball* so it [Diff;Ex;FtC] can only spin in one direction.

Animation Toolbar

The Animation Toolbar is used to make your scene come to life. Any time a scene is changed over time, through drag and drop or manual animations, *the Animation Toolbar* stores that information. It's [Same;Ex;DK] a powerful tool that gets discussed in Chapter 7.

4. Type "Swift 3D" to replace the original text. (*The program* will not crash for those

moxy-laden people who type in a more original piece of text, but keep it [Diff;Ex;DK]on the short side.)

Zoom the camera
Goal: Zoom out on *your text* so it [Diff;Ex;FtC] fits within the Viewport

2. Click and drag on *the Sphere* and position it [Diff;Ex;FtC] just to the left of your text.

3. Slowly drag towards the center of *the sphere* to make it [Diff;Ex;FtC] slightly smaller than it [Diff;Ex;FtCCh] was.
3. Hold the left mouse button down and rotate *the trackball* to see how it [Diff;Ex;FtC] rotates the objects.

Goal: Relocate *your object* exactly where you want it [Same;Ex;FtC]

2. Hold *the left mouse button* down and move it [Same;ex;FtC] to get the Camera Pan cursor to appear.

3. Click and drag on the trackball and position *the spotlight* so that it [Diff;Ex;FtC] shines on your text from the upper right side. Release mouse button.

4. Click on *one of the Default Lights*. It [Diff;Ex;FtC] will become a red wireframe when it [Same;Ex;FtCCh] is selected.

5. Using *the trackball*, position it [Same;Ex;FtC] near the spotlight.

4. Write your Mom *a quick email.* It'll [Diff;Ex;FtC] make her happy. Plus, you've got a moment or two to kill.

Swift 3D is now going to render your animation. What happens is your image will be turned into a whole bunch of shaded triangles. *This process* is what makes Swift 3D so unique because it [Same;Ex;DK] can take *a standard image that looks like bitmap* and turn it [Same;Ex;FtC] into a vectorbased image, thereby saving you some serious time and space.

If your Mom doesn't have email you can sit back and watch *your processor* max out as it [Diff;Ex;FtC] calculates every little curve and angle, *creating a complex wireframe mesh of your image.* The real time-sucker is the fact that it [Diff;Im;DK] has to be done for each frame of your animation. But hey, have you ever tried to create realistic 3D with Flash? We'd be looking at writing a novel to your Mom rather than email.

Totally Text

Text is considered an object. The way Swift 3D turns text into an object is by taking *each character*, giving it [Same;Ex;FtC] some depth (extrudes it [Same;Ex;FtCCh]), and applying a default bevel to the edges. Any font you have installed on your computer can be utilized by Swift 3D to create text objects. Some fonts look better than others when they are endowed with that third dimension.

Smoothness

This slider controls how accurately curves are drawn on your text. If you adjust towards Fine, your curves will smooth out, but the amount of lines it [Same;Ex;DK] takes to render the text increases, and so does your file size. Adjusting towards Coarse will make any curves appear more angular and reduce file size. The general rule of thumb is "leave it [NA;NR;Coll] alone unless you have a good reason not to."

Sizing

Any numerical changes will use the center of the text as a baseline. So if you increase the depth, the front of *the text* will move towards you (if it's [Diff;Ex;FtC] facing the camera) and the back of the text will move away from you.

When you create a piece of text (or import an EPS Image see Chapter 8 for details) you end up with characters that have three types of surfaces. Faces refer to the front and back of the text. If you had *a two dimensional character* it [Diff;Ex;FtC] would have a front and a back face and nothing more Edges are what happens when your text gets depth. They are the surfaces that are perpendicular to the front and back faces.
When you create *a text object*, it [Diff;Ex;FtC] is considered one main object that consists of individual grouped characters, but the characters within the word are considered Children and can be somewhat manipulated independently.

To apply different materials to Individual characters, hold down the CTRL key while dragging and dropping your desired material. You will see *each character* get a bounding box as you drag the material over it [Diff;Ex;DK]. Release the mouse button when you are over your chosen character.

The Hide command will make *your selected object* disappear. It [Diff;Ex;FtC] will still be a part of your scene, but it [Diff;Ex;FtCCh] will remain invisible. One important thing to remember is that for the object to actually be hidden, you must also uncheck the Hidden option under the layout Properties which tells Swift 3D whether or not to show hidden objects
Cone

When you drop *a Cone* into the scene it's [Diff;Ex;DK] just that, a cone. However, a cone is actually a cylinder with a closed end (picky, picky, I know), so when you create *a cone* you have the option of turning it [Same;Ex;DK] into a cylinder by checking the Cylinder option In the Property Tools. With a cylinder you can only control its radius and length, but with a cone you have control over its top radius, bottom radius and length.

Cube

Achieving a cube in Swift 3D is a little different than using the object buttons. Instead, you actually use a text character to create a 3D square and then give *the character* sufficient depth or extrusion to make it [Same;Ex;FtC] a cube.

Grouping and Ungrouping

The Arrange, Group command will take any selected objects and join them to ether to create one object. This allows you to move them around your scene as one unit. It [Same;Ex;DK] also allows you to animate multiple objects together, which can save time over moving each object individually, If you want to break your objects apart after you have grouped them use the Arrange, Ungroup command Aside from the materials that come with Swift 3D, feel free to invent as many as you like and store them in the Materials Palette with the pre-made ones.

Diffuse Light is direct, angle-dependant light. It [Same;Ex;FtC] refers to light from a distinct source that directly strikes an object; for instance, if you hold an object under a reading lamp. In Swift JR diffuse light defines the color of the object, so if you choose green for your diffuse light, you'll have a green object Diffuse light is brightest when the light source is shining directly on the object, or at a 90 degree angle from the surface face. As that angle of incidence decreases (becomes more parallel to the surface) the diffuse light will become more and more dim, thus the angle-dependant stuff mentioned earlier.

Ambient Light refers to uniform reflected light. Referring to our talk about environment vs. background, we used the analogy of the room. Even if *an object* has no direct light on it [Diff;Ex;DK], it's [Diff;Ex;FtCCh] usually illuminated by ambient light since those crazy light rays are always bouncing around and coming at us from every angle. Therefore, the color of the surfaces this ambient light is reflecting off of can have an effect on the color of the ambient lighting. For example, if you place a yellow ball in a blue room, *the ball* will still be yellow, but it [Same;Ex;FtC] will have a slight blue tint from all of that blue ambient lighting.

Highlight is the bright spot *an object* gets when a distinct light source is directed at it [Diff;Ex;DK]. The glossier an object is, the more distinct the spotlight will be. *By changing the color of the highlight,* it's [Diff;Im; FtC] like changing the color of the light source being aimed at the object, as If you had a colored spotlight, In Swift 3D, you can also adjust the size of the highlight on an object via the Materials Editor.

Double illuminate has to do with some complicated things *Swift 3D* does when it [Same;Ex;FtC] determines how to shade each individual polygon in an object. Basically, by creating a material that's double illuminate you are telling Swift 3D to apply that material to both sides of every surface, even if they aren't visible. For example, if you make a cube and place your camera inside that cube, with a double illuminate material you will be able to see the inside surfaces. Without a double illuminate material you would just see black. We feel that double illuminate materials are more versatile so all of the materials that come with Swift 3D are double illuminate.

NOTE: The materials provided with Swift 3D are not editable directly. To create a material based on an existing one, you just make *a copy of the material,* edit it [Same;Ex;FtC] to your liking and save it [Same;Ex;FtCCh] with a new name.

CHAPTER 6: LIGHTING

First, it's [NA;Ex;Coll]crucial *to understand where the lights you are creating and*

manipulating actually reside. Let's talk about that crystal trackball for a minute. It's [NA;Ex;Coll] fairly easy *to envision point lights and spot lights attached to that sphere*, shining into the center of your scene. The difficult part is realizing how big that sphere is The rule is that your 'sphere of lights" is as close to the objects as possible, while still encompassing all of the objects within the scene. In plaintalk - the bigger your scene, the bigger the sphere of lights,

What this means is that as you create and move objects around in your scene, the lighting scheme will interact with the objects producing different effects. For this reason it's [NA;Ex;Coll] a good idea *to lay off any detailed lighting designs until you' ve got your scene fairly stable.*

As we talked about in Chapter 2, Point Lights are like light bulbs and spot lights are like flashlights. So if you are looking for specific effects, choose your weapon carefully. Point lights cast more light overall. It's [NA;Ex;Coll] easy *to flood out your objects and lose the 3D effect with too many regular lights placed towards the front of your scene.* Spot Lights don't cast much light and should be used mostly to create highlights or bright points in your scene. Spot lights are also very sensitive to the direction they point, so when we talk about placing lights directly into the scene, spot lights are the most effective tools.

Adding and Subtracting lights is a very simple process with the buttons, When *a light* is selected it [Same;Ex;FtC] becomes a wireframe, and only selected lights can be removed with the Subtract Light button. Once *a light* is selected, you can move it [Diff;Ex;FtC] around the scene by clicking and dragging anywhere on the lighting trackball. When *a light* moves to the back of the scene it [Same;Ex;DK] gets smaller. You can only move one light at a time.

Lighting Properties

When you have a light selected, you can view that light's properties in the Property Toolbox. With a Point Light you have control over *the color of the light* and whether it [Diff;Ex;DK] is active or not. By double clicking on the Color Window you access the Color Palette that allows you to choose from some pre-made color or create your own, For a more in-depth explanation of color creation see Chapter 5 on materials.

4. You can now treat the light as a full-fledged object, using the same mouse maneuvers you would use to position objects. The best way to aim a light placed into the scene is with the object Trackball. Although you can *resize * the light with the scaling tool*, it [Diff;Im;DK] has no effect on the amount of light cast. (Read: don't bother.)
*resizing is the referent

Background and Ambient Light Color

To be honest, manipulating the Background is of limited use to you since the majority of Images created in Swift 3D will be imported into a SWF editor for further manipulation. When you bring *an exported SWF file* into Flash, the background does not come along with it [Diff;Ex;DK], only the objects. If, however, you play your Swift 3D creations in the Flash Player, the background is respected and will appear in your scene.

To be honest, manipulating the Background is of limited use to you since the majority of Images created in Swift 3D will be imported into a SWF editor for further manipulation. When you bring an exported SWF file Into Flash, the background does not come along with it, only the objects. If, however, you play your Swift 3D creations in the Flash Player, the background is respected and will appear in your scene. As for _Ambient Light Color,_ it's [Diff;Ex;FtC] a subtle effect. For the average user its scope is limited. The main reason you might alter this setting is to lend a slight tint of color to gray or light colored objects, or to lighten or darken your objects overall.

Before we buck our brains into _the gripping world of animation,_ let's chat a lad _about applicability._ The basic question you have to ask yourself is, "is it [Diff;Im;DK] worth it [Diff;Im;DK]?" To determine the answer, you should take a look at the big picture here, Swift 3D has solved the confounding problem of creating quality D images in a vector-based format. Swift 31) also has some great animation tools that allow you to create dynamic scenes. But Swift 3D is currently just a stepping stone for images on their way to stardom in another vector-based 2D presentation.

Keyframes

Understanding Keyframes and how they work is critical to creating successful animations. Basically, keyframes represent points in an animation where you can make important changes take place. In between two keyframes, Swift 3D will make all the decisions on how your scene changes using its tweening function. The keyframes simply give _Swift 3D_ reference points to work from while it [Diff;Ex;FtC] decides how to create the transitional frames.

Each keyframe has a Start Animation and a Stop Animation control. Within one Frame the controls are set to a value of one so there is no pause. Adjusting these controls allows you to stop one aspect of tin animation for as long as you want. For example, you may want a spinning text object to pause mid-spin so it's [NA;Ex;Coll] easier _to read._ We've used the universal 'Green equals Go. Red equals Stop" color scheme to make it [NA;Ex;Coll] obvious _when an object is going to move, rotate or change size,_ and when it [Same;Ex;FtC] is going to stop doing those things.

The Frames Per Second Control allows you to adjust how smooth your animation will be upon export. The more frames that get viewed each second, the more continuous your animation will be, and visa versa. _The default of 12_ has proven itself to be the most effective for use on the Web, and when you import Swift 3D files Into Flash it [Same;Ex;DK] overrides any changes and reverts back to the default of 12 frames per second.

NOTE: This command only deletes the selected objects. All other objects that have animations are not deleted. Often times _the camera will get animated_ * without you realizing it [Diff;Im;DK]. so make sure you check that if you are trying to get rid of all animations. *animation

When you right click on any of your keyframes there is an option to control Properties. With this feature you can adjust how _the animation_ behaves as it's [Same;Ex;FtC] played. The three main adjustments you can make are Tension, Continuity and Bias These features are designed to influence how paths and scaling behave, not spins.

Tension controls how smoothly the animation flows through the keyframe. With less tension the animation will flow more naturally through each keyframe. With more tension *the animation* will have a more abrupt feeling as it [Same;Ex;DK] moves through each keyframe.

Continuity is a similar control only it [Same;Ex;DK] allows you to force the animation past its designated keyframe. For example, if you apply a high continuity to a scaling animation where the object is getting smaller then larger, *the object* will get even smaller than the keyframe, then bounce to the keyframe size, then get smaller again before it [Same;Ex;DK] finally increases in size to the next keyframe.

Bias allows you to control whether *an animation* flows equally through a keyframe or whether it [Same;Ex;DK] favors one side or the other. For example, a curved path animation would be more heavily curved on the side you adjust the bias towards.

2. Make your animation as short as you can while still maintaining decent quality. The more_frames_to your animation, the longer it [NA;NR;Coll] will take to render and the larger your file size will be.

3. Make sure that you are in the proper frame before you make any changes to your scene. It'[NA;Ex;Coll]s easy *to play your animation*, stop it [Same;Ex;FtC], and change something you don't like. Whatever frame you stop your movie on is going to be the frame that gets changed, and if you're using tweening functions to create your animations, you can goof up your work pretty quickly.

4. Select *the Spot Light* and move it [Same;Ex;FtC] where you want in the scene. Right click and drag will move it [Same;Ex;DK] forward or back. Use the Object Trackball to aim the Spot light.

Swift 3D allows you to animate any of your camera views by using the Pan function. It [Same;Ex;DK] also provides one camera view that moves from a pivot point which can lead to very realistic "look around" animations.

The big difference between these two animated camera options is that Swift 3D's Standard Cameras are designed so that when you pan them they are actually being repositioned on a two dimensional plane rather than actually panning. The Rotating Camera acts like *a real camera* would as it [Same;Ex;FtC] pans - moving around from a fixed point. The Rotating Camera will create more realistic animations as if you were moving your head around and looking at various objects in your scene.

While you are creating your masterpiece In Swift 3D, the file being generated is In the T3D format- So if you want to insure you have an editable back-up of your work, or if you need to save an unfinished project, the Save As and Save commands work just like any other program. When you are totally done designing your image or animation, the Export command creates a vector file that Flash uses (SWF), which can be imported directly into Flash. Once *the file* is exported, you can no longer manipulate it [Diff;EX;FtC] using Swift 3D so you want to also do a Save As before you quit Swift 3D. This will give you both the exported file and a T3D file for future editing.
When you bring d 3D Studio model into Swift 3D, our program reads the objects,

materials and lighting schemes that were created in the original 3DS scene. This process is not really a simple Importation of an object since 3DS files have all this associated baggage that Swift 3D has to deal with. In other words, rather than importing a *3DS file*, you are opening it [Same;Ex;FtC] Into Swift 3D.

Another thing that comes along with these 3DS files is the rotating camera function. In Swift 3D *the rotating camera* pivots around itself, sort of like it's [Same;Ex;FtC] on a tripod. In 3DS you can designate a camera's target point so that the camera rotates around a fixed point. Swift 3D respects that functionality and you end up with another option in the Camera View menu that gives you access to that imported camera view.

When you bring your 3D Studio model Into Swift 3D, *the model* could appear as a grouped object or an ungrouped object. It [NA;NR;Coll] depends completely on whether it [Same;Ex;DK] was grouped upon export in 3D Studio, so your best bet is to select the object and look in the Arrange menu to see what *your model's status* is. If it's [Same;Ex;FtC] ungrouped and you would like it [Same;Ex;FtCCh] to become one object, do an Edit, Select All and then an Arrange, Group and you'll be in business.

Encapsulated PostScript (EPS) and Adobe Illustrator (AI) files

Swift 3D will take *almost any vector EPS image* and turn it [Same;Ex;FtC] into a three dimensional object, while retaining its original 2D shape. You will lose any previously applied colors but those can be controlled from within the Swift 3D program, So vector-based drawing programs that export to the EPS file format can now function as preliminary design tools for later use in Swift 3D, In addition, Swift 3D supports the importation of Adobe Illustrator (AI) files.

The only complication you may have is if you have imported *a bitmapped image* into your drawing program and then exported it [Same;Ex;DK] as an EPS or AI image. Swift ID does not recognize these bitmaps, so after importing some cool artwork, you may find yourself screaming in horror or at what our program has done to your image.

NOTE: If *your image* is made up of more than one shape, it [Same;Ex;DK] will arrive into *Swift 3D's Interface* as a grouped object. You will not be able to change its properties until you have ungrouped it.[Diff;Ex;DK] You will then have access to the same properties you would have if the original EPS Image was a single object.

When you go to export your image you will be faced with several options and formats. Generally, there will be a standard that you use when exporting, but it's [NA;Ex;Coll] important *to know what all these functions are so if you have any special needs you'll know how to meet them.*

Exporting your Image as both outlines and fills will create two separate Images layered on top of one another. One will be the lines between all of the surfaces of your objects and the other will be the fill patterns that give your objects their solidity. The only advantage this export formal has over just fills is that when you bring your image into Flash you can keep the two separate and create some cool effects. It's [NA;NR;Coll] your image - It's [NA;NR;Coll] your decision. If you use this formal and want to keep the lines and fills together as one image, make sure you group them together as one object when you import them into Flash.

There is a limitation to the complexity of objects you can export from Swift 3D. It's [NA;Ex;Coll] safe *to assume that anything you design within the program itself will have no problem being exported.* It's [NA;Ex;Coll] when you start bringing in very complex models from outside sources like 3D Studio *that you might exceed our current polygon limit.* If this happens, you will be greeted with a kind message from our engineers saying that your image is just too overwhelming for our current version of Swift 3D. And trust us, we'll be working like dogs to get you *a new version* that can handle absolutely everything you throw at it [Diff;Ex;DK].

Mesh Shading is the best you can get. It [Same;Ex;FtC] shades each polygon in the mesh producing greater accuracy. *Your exported image* will basically look exactly as it [Same;Ex;FtC] did in your Viewport within Swift 3D You will retain all of the lighting, reflections and details of your image, but you'll also end up with a file that's pretty sizeable. *Best shading - Largest file size Outlines* is the simplest of export options. It [Same;Ex;DK] just shows the hard edges of your objects, or the places where distinctly different planes adjoin. If you are using the flat shading export format the outlines can help considerably in defining your objects without adding much file size.

One way or the other, you're most likely going to be importing *your image* into a Flash editor. What you do with it [Diff;Ex;DK] from there will decide how you want to export it [Diff;Ex;FtCCh] from Swift 3D. If your Flash file is heading for the Web, you can pretty much bet that *Area Shaded* is going to be the most realistic export format for you. It [Same;Ex;DK] retains fairly good quality while shrinking your file size down considerably. Just to give you an idea of how much your file size will be reduced, you're looking at a bare minimum of two times smaller than a Mesh Shaded export.

Both Outlines and Entire Mesh have their uses, but realize that if you're going for a realistic looking 3D Image, the lines just don't get you there. If gradients are not a priority for you, Average Fill provides a nice 3D effect and the file size will be a bit smaller than an Area Shaded export. And as far as *Flat Filled* is concerned, don't waste your time unless you're going to be using it [Diff;Ex;DK] in conjunction with outlines.

As long as you save the imported SWF as a FLA file, you should never have the polygon dropout issue bother you again (for that file), keep in mind that if you export *the file* as a movie from Flash, and then re-import it [Same;Ex;DK], you will see *the polygon dropout problem* again. Sorry for the bad news, but at least we have a temporary solution for it [Diff;Ex;DK] until Macromedia solves the issue.

Finally, there is *one more problem* between Swift 3D and Flash that is yet unresolved. It [Diff;Ex;DK] too involves dropped polygons On occasion, some of the polygons will be missing from the scene after import into Flash. This problem occurs in both Flash 3 and Flash 4 and is different from the one discussed above. In this case, you may experience *a few polygons missing* within a few frames in a 10-20 frame average animation. It [Diff;Ex;DK] could be more if your model is very complex or you may never see it [Diff;Ex;FtCCh] in a given animation. This problem occurs most often with Mesh Shaded scenes and very rarely with Area Shaded scenes and never (we think) with Flat Filled scenes, we are currently working with Macromedia to remedy this problem so check our Customer Only Web site for up-to-date information regarding a fix for the problem.

APPENDIX 2

ABBREVIATIONS FOR THE ANNOTATION SCHEME

Syntactic Function

Diff	Different
NA	Non-Applicable

Type of antecedent

Ex	Explicit
Im	Implicit
NR	Non-Referential

Processing Strategy

Coll	Collocation
FtC	First Candidate Search
FtCCh	First Candidate Chain
DK	Discourse Knowledge

APPENDIX 3

E-MAILS

De:	nvandenheuvel@lhsl.com
Enviado em:	Terça-feira,22 de Fevereiro de 2000 09:30
Para:	rosar@cce.ufsc.br
Assunto:	Power Translator Pro

Prezada Acacia,

Thank you for your interest in our translation product. PTP6.4.3 is available in most computer software stores throughout Brazil. As the product is manufactured by our representative in Brazil, MSDMultimidia, you might best contact them directly.

Once you have completed your analysis, please send the results to me directly, and I will personally see that you receive a token of our appreciate, such as a free upgrade to the next version or another fine L&H product.

Best of luck in your studies.

Nicole A. VandenHeuvel
Channel Sales Manager
Central/South America

De:	marciot@msdmultimidia.com.br
Enviado em:	Quarta-feira,01 de Março de 2000 14:09
Para:	acacia@fln-bma.yazigi.com.br
Assunto:	Re: Power Translator Pro

Acacia,

Não estamos envolvidos em nenhum projeto deste tipo porque não somos os desenvolvedores dos produtos e, portanto, não estamos capacitados tecnicamente para esclarecer dúvidas ou implementar melhorias nos códigos utilizados na elaboração e compilação dos mesmos. A MSD Multimídia apenas licencia e distribui, em território nacional, títulos em CD-ROM oriundos de outros países. Somente a própria LHs poderia estar lhe oferecendo informações deste tipo.

Atenciosamente,

Márcio Torres – Suporte Técnico MSD Multimídia e Informática.
(0xx21) 533-3919.

De: customer.service@lhsl.com
Enviado em: Segunda-feira,6 de Novembro de 2000 21:05
Para: acacia@fln-bma.yazigi.com.br
Assunto: Re: information

Dear Acacia,

Thank you for inquiring about our Machine Translation technology. To answer a few of your questions, the translation process is executed by the L&H Barcelona Server Engine. The Lexical rules within our program addresses the question of ambiguity. Because L&H's exchange of research data is regarded as trade information, L&H provides general information about its Technologies on the website. Hope this information is helpful, feel free to visit our website at www.lhsl.com.

Thank you,

Dawn
L&H Technical Support

Acácia Rosar <acacia@fln-bma.yazigi.com.br> on 10/30/2000 08:50:04 PM
Please respond to <acacia@fln-bma.yazigi.com.br>
To:<customer.service@lhsl.com>
cc:
Subject:information

I have Global Link Power Translator Pro Version 6.4 and I'd like some
information. I'm an MA student at the Federal University of Santa Catarina - Brazil. I'm doing a research in computational linguistics more precisely about MACHINE TRANSLATION, analyzing the main problems and solutions to the English pronoun IT in terms translating. I'd like to know whether there is any form of cooperation with research initiatives presently active in Lernout & Hauspie, such as information exchange, technical support, particularly in discussing solutions regarding to the problems that occur while translating the pronoun IT, especially into Portuguese. was this program developed within a corpus-based approach? What are the main strategies used to translate the pronoun "IT" into Portuguese? How does the machine deal with the question of ambiguity? Waiting to hear from you soon.

Acácia Rosar