

Chapter 6

Many of us grow up with the idea that mistakes are bad, linking our self-esteem with continued success. We become afraid of making mistakes. So in order to achieve success, we tend to avoid areas that may lie outside the apparent realm of our natural talent. In this perverse equation, the secret of success becomes avoiding failure, leaving much of our potential untapped. In order to reach our full potential to learn, we must accept and then transform anxiety and fear, relentlessly seeking accurate information on our performance. What used to be perceived as criticism now becomes a gift for constructive growth.

Michael Gelb, Tony Buzan

6 Towards Cross-Cultural Information Retrieval

6.1 Introduction

In acquiring an actual cross-cultural retrieval capability it will require more than learning new approaches of communicating information. It will require the learning of new approaches of indexing and describing information.

Transforming one's awareness (a sense of one's personal or collective identity that includes your attitudes, beliefs and sensitivity) makes it viable to naturally accommodate cross-cultural factors when encountering people of dissimilar cultures. Instead of applying rules of "culturally-appropriate behaviour" one should rather speak and behave naturally.

This chapter intends to identify those cultural factors by categorising the errors experienced in the analysis in chapter five, and highlighting the most important issues to be addressed in future research. If this is successfully dealt with, it is possible that cross-language information retrieval could give rise to cross-cultural information retrieval.

6.2 Categorising the problems identified in the error analysis

Through a detailed analysis of errors in Chapter 5, two main problems (each with its own sub-categories) were identified. They are dictionary problems and translation problems. A summary of these problems and their sub-categories are presented in Figure 6.1:

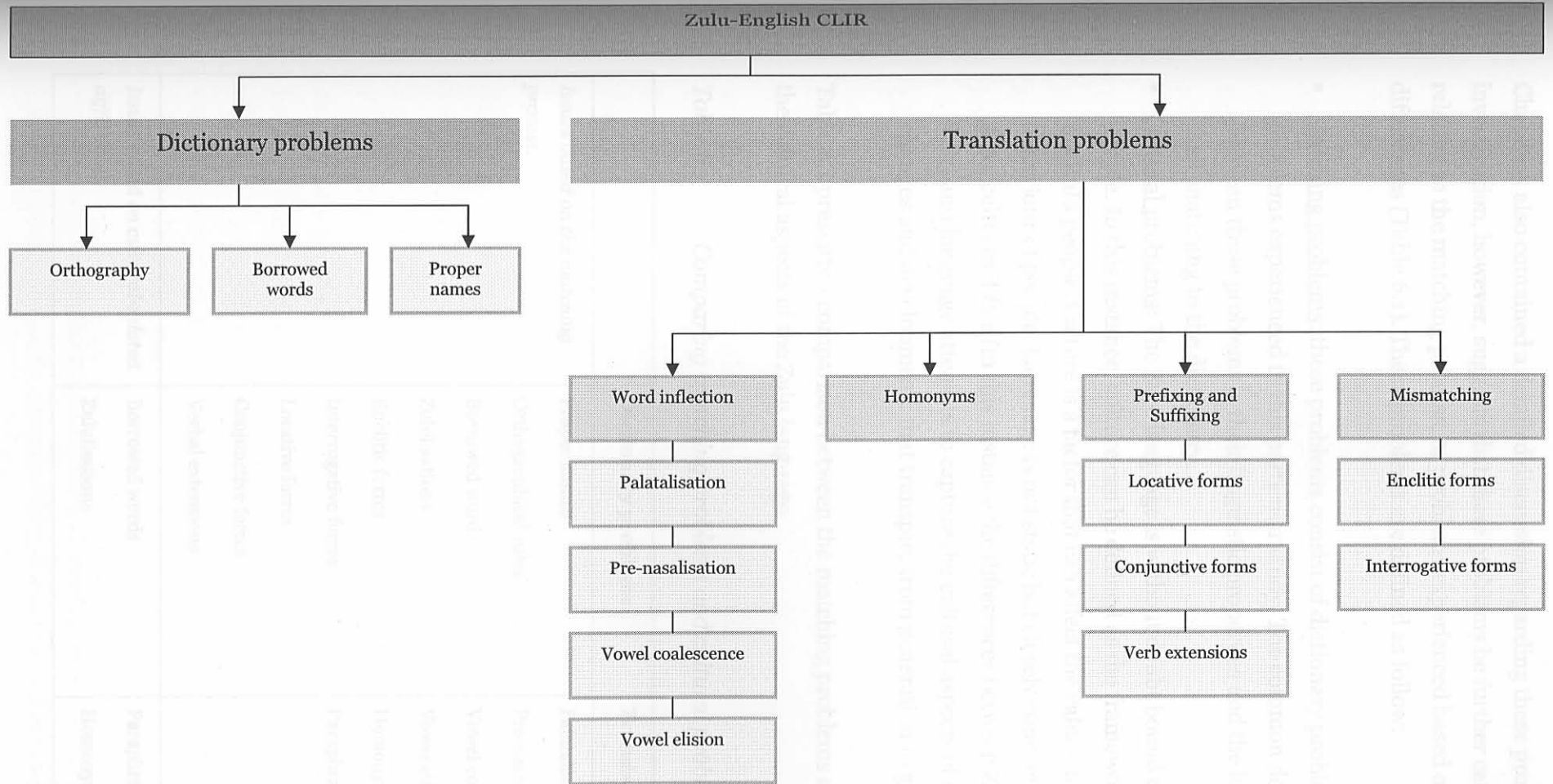


Figure 6.1 A schematic representation of the analysis of errors

Chapter 5 also contained a detailed discussion regarding these problems. Further investigation, however, suggests that these problems be further categorised as either relating to the matching process, or problems experienced based on cultural differences (Table 6.1). These problems are defined as follow:

- **Matching problems:** these problems consist of dictionary problems and translation problems experienced in this particular study. The common denominator between these problems is their linguistic properties and the impact it has on their matching to the dictionary.
- **Cultural problems:** The Zulu language is without doubt bound to the culture of its people. In this instance, culture can be defined as the framework for the lives of the Zulu people. Culture is a factor that can affect the values, attitudes and behaviour of people. Language is not static but closely connected with culture. Also, cultures differ (in this instance the differences between Zulu and English is vast) and language attempts to capture the cultural aspects of the society's changes and developments that transpire from generation to generation.

Table 6.1 presents a comparison between the matching problems and those related to the cultural aspects of the Zulu language.

Table 6.1 Comparing matching problems and cultural problems

	Dictionary problems	Translation problems
<i>Issues based on the matching process.</i>	Proper names Orthographical rules Borrowed words Zululisations Enclitic forms Interrogative forms Locative forms Conjunctive forms Verbal extensions	Palatalisations Pre-nasalisation Vowel coalescence Vowel elision Homonyms Paraphrasing
<i>Issues based on cultural-related differences.</i>	Borrowed words Zululisations	Paraphrasing Homonyms

From the matrix, one sees that some of the problems are not only matching related, but also cultural in nature. These issues will now be discussed.

6.2.1 Issues based on the matching process

The main cause for the following problems (both from a dictionary point of view as well as from a translation point of view) is that the words change so much in form that it becomes difficult and sometimes even impossible to match them to the original root words in the dictionary.

The dictionary problems stem from the orthographical rules applied to words that make it difficult for matching, while translation problems are mainly because of grammatical changes to words when certain occurrences are noted in specific instances.

On a matching level, it is evident that the Zulu language borrows extensively from English, Afrikaans and the Khoisan language. These borrowed words may take two different forms, namely words that are **adapted to Zulu pronunciation** and words that remain unchanged with only **class prefixes added to conform to the grammatical system** (Section 5.3.2). Even though the pronunciations are similar, the spelling of these Zululised words is vastly dissimilar from English words. This brought about that most of the borrowed words did not match any dictionary entries. Proper names are generally not translatable, which cause matching problems. This is because most proper names are not found in the dictionary. There are a few exceptions though, where words are found as Zululised forms (*iFulansi* for France in Table A15.1; *iYurobhu* for Europe in Table A21.1; and *eJalimane* for Germany in Table A26.1).

From a dictionary point of view two additional matching-related problems are the enclitic form (*-nje*) and the interrogative form (*-phi?* or *-ni?*). These words are normally added as suffixes to verbs or adverbs, but they are sometimes matched to dictionary entries. However, they do have a clear syntactical meaning. In the empirical experiments done, it was clear that these word forms were incorrectly matched to dictionary entries. This had a significant impact on the retrieval results.

Furthermore, the prefixing and suffixing of words (Section 5.4.4) also have a definite impact on matching results, since matching the words in their locative form, their

conjunctive form or as verbal extensions are problematic to achieve accurate results. This happens because these word forms are not words themselves, but they are mostly mismatched to parts of dictionary entries.

There are several orthographical rules that the Zulu grammar abides to, and which are indicated in Table 5.1. When some of these rules are enforced in the written form, certain problems arise. In some instances, it was difficult to look up the words in the dictionary, as they could not be found under the assumed entry. This was the case with a word like *-bhala*, which was actually found under the dictionary entry *-bala*. In pronouncing the word *-bhala*, one would never be able to distinguish the two different letters (b).

In the instance of paraphrasing, it seems as if the dictionary translation attempts to describe each of the words in phrases of three to four words. For instance temperature (the process of becoming hot); air pressure (the influence of the wind); and rainfall (the coming of wetness to the sky). These examples are indicated in Table A3.2. When the mother tongue translators were queried as to why the dictionary and mother tongue translation differ so much in the paraphrased word form, their reply was: “we do not really speak like that”.

These forms of paraphrasing are the result of directly attempting to match Zulu words to dictionary entries. These are just a few examples that demonstrate where no matches were made (in the instance of *nesimo* and *lokushisa*) and where mistranslated matches were made. In the instance of *nesimo*, the noun stem is actually *-mo*, and therefore no match was made. This also accounts for *lokushisa*, which is found in the dictionary found beneath *ukushisa*.

Zulu is a highly inflectional language (30% of the errors were because of word inflection), and matching the words in the running text to the dictionary entries are often problematic. This occurs because of the changes made to the syntactic function of the word, thus being inflected for the singular and plural, for future and past tenses, and other tenses. The different palatalised and pre-nasalised forms are indicated and discussed in detail in Section 5.4.2.1 and Section 5.4.2.2. An applicable example is where the letter **b** in *-sebenza* changes to **tsh** to form *setshenziswa* (the causative *-is-* and passive *-w-* are added as verbal extensions).

Other word inflection forms described in Section 5.4.2, are vowel coalescence and vowel elision. Vowel coalescence occurs when the basic Zulu vowel *a* (usually found

in the instrumental *nga-* and the adverbial formative *na-*) coalesce to *nge-* or *ngo-* when followed by the vowels *i* or *o*. The latter occurs when two word groups are combined to form a new word. This occurs apart from the grammatical elision of initial vowels (when forming vocative interjections or after adverbial formations). Several other examples are presented in Section 5.4.2.3 and Section 5.4.2.4. The negative impact of these word inflections are evident in the results that were obtained, mainly because it was difficult to correctly identify the instances when words coalesce, or when palatalisation takes place. Because palatalisation occurs automatically, the matching process could not identify the source language words (appearing in their inflected forms). This is because the words that these inflected forms are being matched to, appear in their base forms as they appear in the dictionary.

6.2.2 Issues based on cultural differences

There is a vast difference between the Zulu and the English culture, which makes CLIR difficult. The Zulus have a different way of performing their daily tasks, and in expressing themselves. Through pictures, metaphors and idiomatic sentences they provide a more detailed explanation of their surroundings and feelings. This is not so for English, and because of these differences, some instances occurred where words did not match dictionary entries. Furthermore, the Zulu language frequently makes use of tone to indicate the meanings of words. Sometimes the exact same letter in words (spoken in a high or low tone) indicates different meanings. When the Zulu language is lacking a word for a certain concept in English, a new word is immediately created. This is evident in several of the Zululised words (Section 5.3.2). Not only is the paraphrasing of words difficult to match on a conceptual level, but the cultural aspects should also be taken into consideration in the matching process. This is mainly due to the richness of the Zulu language in describing words and actions. Paraphrasing is commonly applied to describe words or phrases in technical or scientific terms. Examples of paraphrasing are found in Table A1.1 (pesticides = *amahkemikheli abulala zonke izifo ezinengozi*) and Table A21.1 (synagogues = *kwezindlu zesonto zamaJuda*). More examples are discussed in Section 5.4.1. In the example presented, the word synagogue is paraphrased to “the dwelling place of Sundays of the Jews”, but in Table A21.1 the mother tongue translation actually translates the word to *amasinagoge*, which is a borrowed form (Zululisation) of synagogue.

In Table 4.1 the word ‘tour’ occurs. The dictionary does not match the word, but leaves it as a proper name, and add a Zulu class prefix. The mother tongue translation though, translates the word to “the competition of the journey that is endless”. Here the translator might have heard about the Tour de France, and placed it in his/her own context of an endless journey (possible because 3350km on a bicycle seems like a journey that would never end). If the word is left in its original form (‘tour’), acceptable results can be obtained when the n-gram technique matches the word to its base form (and not the paraphrased form).

Although there is not really a Zulu word for corruption (Table A15.1), the mother tongue translator indicated that the word *ngenkohlakalo* is generally used for corruption in a sense of “events or happenings of a criminal nature”.

Another example where cultural aspects play a role in moving towards cultural CLIR, is where the word “green power” (Table A33.1) is mistranslated in the mother tongue translation. The mother tongue translation has nothing to do with electrical power, but rather power in the sense of strength (and then it must also be green in colour).

In Table A34.1 the query concerns Mad Cow disease and retrieving documents related to this sickness. In the mother tongue translation, however, the term “mad cow” is translated to “the disease of the hoofs”. This is primarily because Zulus are aware of the disease, and they know animals with hoofs are the cause of the sickness – therefore they provide the culturally correct translation. The dictionary translation again only applies the rule for borrowed words and adds the class 6 prefix to obtain “i-Mad Cow disease”.

Through paraphrasing, it was found that the association between two classes did not always contribute to the user’s (in this instance the translator) understanding of the text. This is one of the main problems of paraphrasing – it is a *mechanical* process of exchanging words and phrases for synonyms. In effect, it discourages careful consideration of the meaning of the text itself. Consequently, the translator may interpret the text differently and as a result make small changes in the written text. In doing this, the translator actually changes the meaning of the original text itself. This causes the amount of text paraphrased being too general or unfocused, which was also not always effectively translated or matched on a conceptual level.

In the instance of homonyms; the Zulu language is rich in the several different meanings it has for the same word. The deciding factor that determines the meaning,

is the tone applied in speaking the word, as well as the class prefix added to the word. It is difficult for a machine to recognize where tone is indicated in a word, since the tone is created by means of an acute symbol in most instances. For example: the word *-zwe* are found in several instances, and could translate to either 'nation' or "rapidly spreading brain disease". The same happens with *-ngena* in Table A6.2, where the word could either mean "the milking of a cow", "the purpose of taking a wife", "to enter the new year" or "to know intimately". Another interesting example is found in Table 10.2, where the word *-su* could translate to 'plan', 'method', 'scheme', or "rings on a cow's horn indicating the succession of birth" (This last translation is also a good example of paraphrasing).

In Figure 6.2, the issues (above discussed) relating to matching problems are compared to the culture-related issues. Interestingly enough, the 101 culture-related issues were experienced in only two instances (paraphrasing and homonyms), while the 109 matching related issues were experienced in nine instances (proper names, palatalisations, pre-nasalisation, vowel-coalescence, vowel elision, locatives, verb extensions, the enclitic and interrogative forms). The two instances (borrowed words and Zululisations) where the matching and cultural issues overlapped occurred 48 times.

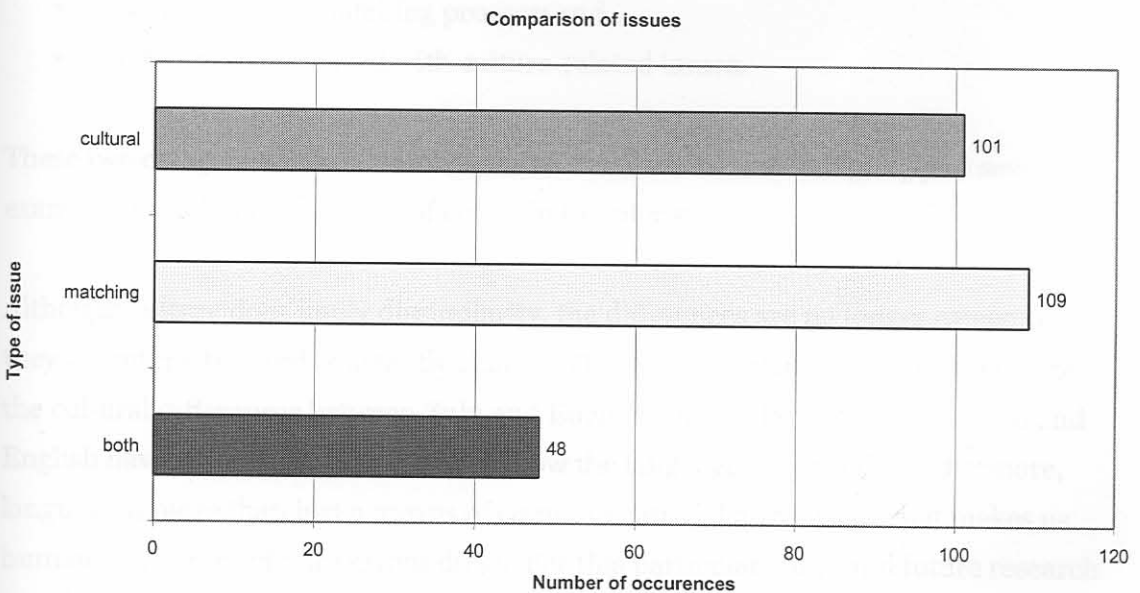


Figure 6.2 *A comparison of culture-related issues and matching issues.*

The reverse process that was tested in Cosijn et al's study (2002a, 2002b, 2002c, 2002d) may solve several of the matching problems experienced (as indicated in

Section 6.1.1 and Figure 6.2). However, the cultural issues need to be addressed in a different manner.

Most of the above mentioned problems encountered in the empirical experiments could be accounted to the morphological and syntactic structure of the Zulu language. It is clear that certain discrepancies exist in the comparison of mother-tongue translations and dictionary translations. The reason for this is mainly because of the contextual element associated with mother tongue speakers, and because dictionary translations are much more of a machine approach. These examples in itself provide a framework for in-depth research into cultural translation and the correct interpretation of texts, as a lot of meaning is lost through the self-interpretation of texts without placing it in a specific context.

6.3 Chapter synopsis

The concept of Cross-Cultural Information Retrieval was introduced in this chapter, and a short definition of what it encompasses was provided. Furthermore, this chapter also provided a concise summary of dictionary and translation problems (identified in Chapter 5) in terms of two new categories to which they could belong:

- problems in the matching process; and
- problems experienced with culture-related issues.

These two categories were identified, and a detailed discussion highlighted several examples found in the analysis of errors in Chapter 5.

Although culture does imply dissimilarity, the differences are no longer categorical; they are interactive and constantly change. This chapter attempted to acknowledge the cultural differences between Zulu and English. Dissimilar cultures like Zulu and English have various interpretations of how the language is applied. Furthermore, language is more than just a means of communication. Language is what makes us human, and aware of our surroundings. For this particular study and future research in CCIR, it is critical to be aware of Zulu behaviour, their beliefs and their values (stemming from personal and professional culture). Culture is the soul of the Zulu people. Through language, information can be passed on, one can learn from and connect with others, form and cultivate relationships, analyze, abstract and evaluate facts and concepts.

Language has several faces and layers, like the people who use it to communicate. An infinite number of possible grammatical sentences can be built with any linguistic inventory. Through the analysis of errors in Chapter 5 and Section 6.2 regarding culture-related issues, several examples indicated how the Zulu language is utilised to analyze thoughts, inform, dream, express emotion, convince, analyse and confuse. The list of linguistic applications is endless. Also, the richness of the Zulu language proves how it evolved over time, and how it became part of the culture that defines it today. This is evident in the rhetorical speech, shifts in meaning, borrowed lexicon and structures, connotations, denotations, alliterations, metaphors, satire, parody, irony, registers, syntax and semantics.

The discussion regarding the cultural aspects will be used specifically to answer the research questions posed in Chapter 1.

Effort only fully releases its reward after a person refuses to quit

Napoleon Hill