surrounding context. In addition, the assessment carried out in the current research is on the lexical and syntactic level.

Test suites are very important in MT development and enhancement. Rules in an MT system may be "too strict or too general to apply correctly in all circumstances" (Doug Arnold, et al., 1994: 176). Test suites help developers detect such errors easily. They can then modify the rules and run the test suite again to measure improvements.

I would like to mention some facts about the test suite I constructed in this thesis. First, it is useful only in testing legal errors since the sentences are extracted from legal contracts. It is not meant to measure translation performance in general. Second, it is designed for English-Arabic translation. Third, the purpose of test examples is not to compare *Google Translate* with other MT systems as this could be done in further research. The aim is to explore linguistic and legal problematic areas surmounted by *Google Translate*.

It must be pointed out that this evaluation, like any other one, is not, nor can it possibly be a comprehensive one. A complete assessment covering all aspects of lexical and syntactic errors made by the system requires many extended volumes, which is beyond the purpose of this research. The emphasis in the current research has been basically on salient drawbacks on both levels which hinder the overall process of translating legal texts by *Google Translate.*

4.5 Qualitative Assessment

In this section, the researcher opts for categorizing the problems which prevail during the assessment. Such categorization could serve the system developers as well as other researchers who aspire for developing the output of *Google Translate*.