# Assessment tools for the training of interpreters: the case of multilingual grammar evaluation

**José L. Martínez-Fernández,** and **José C. González**

DAEDALUS S. A., Centro de Empresas “La Arboleda”, Ctra. N-III Km. 7’300, 28031, Madrid, Spain
{jmartinez,jgonzalez}@daedalus.es

The work described in this paper has been developed in the framework of the ASSESSTI project (Assessment Tools for the Training of "Interpreters") funded by the Leonardo EU programme. The main goal of ASSESSTI is the development of a platform for the support of the learning process of interpreters in four different European languages: Bulgarian, Czech, Dutch and English. This document describes one particular aspect of the ASSESSTI platform: the tool for the semi-automatic assessment of the translations produced by human interpreters during their learning process.

**Keywords** human interpreters; training; linguistic checking; multilinguality

## 1. Introduction

The main objective of the project is to support the learning process of interpreters in a multilingual environment. The ASSESSTI learning platform provides a web-based learning environment and a toolbox for the training and evaluation of interpreters. The platform makes it possible for professional or student interpreters to run learning self-assessment tests. It is also useful for teachers of translation/interpreting, as the platform provides an automatic and detailed evaluation of the progress achieved by students. The platform offers three different management roles: User management, Course Management and Website Administration. The self-assessment functionality provided by the platform takes into account a range of features describing the performance of interpreters, such as:

- *Thesaurus-based features*: Compare source text and target-interpreted text using a predefined multilingual thesaurus. Roughly speaking, this thesaurus holds possible translations of source words into target words. This allows for a validation of the vocabulary used by the interpreter.
- *Acoustic features*: Elements like pauses, hesitations, repetitions, pitch and loudness are captured and analyzed. Speech segments are indexed, tagged and grouped by similarity in order to provide track averaging and quantification.
- *Grammatical features*: Starting from the analysis of linguistic structures and combinations commonly used in the target language by native speakers, these are compared with the same linguistic features captured for the interpreters’ output transcripts. This behaviour gives the module the name of 'detector of non-typical phrases'.

The work described in this paper is centred in the analysis of the grammatical features, covering the development of the non-typical phrases detection component.

## 2. Architecture for the grammatical assessment module

The developed tool is in charge of the semi-automatic assessment of the grammaticality of a translated text. To be able to assess grammatical quality, the first thing needed is the knowledge about how grammatical structures are used in the target language. One suitable way for obtaining this knowledge is the use of corpora. For this reason, the tool being developed allows for the statistical analysis of linguistically analyzed texts. The measured statistics are configurable, including, among others, the different combinations of syntactic structures inside the texts. The result provided by this statistical analysis is formed by a set of grammatical patterns that the translated text must match to be considered as valid. In fact, a graduated scale with three possible values (weak/medium/strong) has been defined. For the development of the tool, a modular design was considered, in order to allow flexibility and scalability. Figure 1 shows the different modules in which the software tool has been divided.

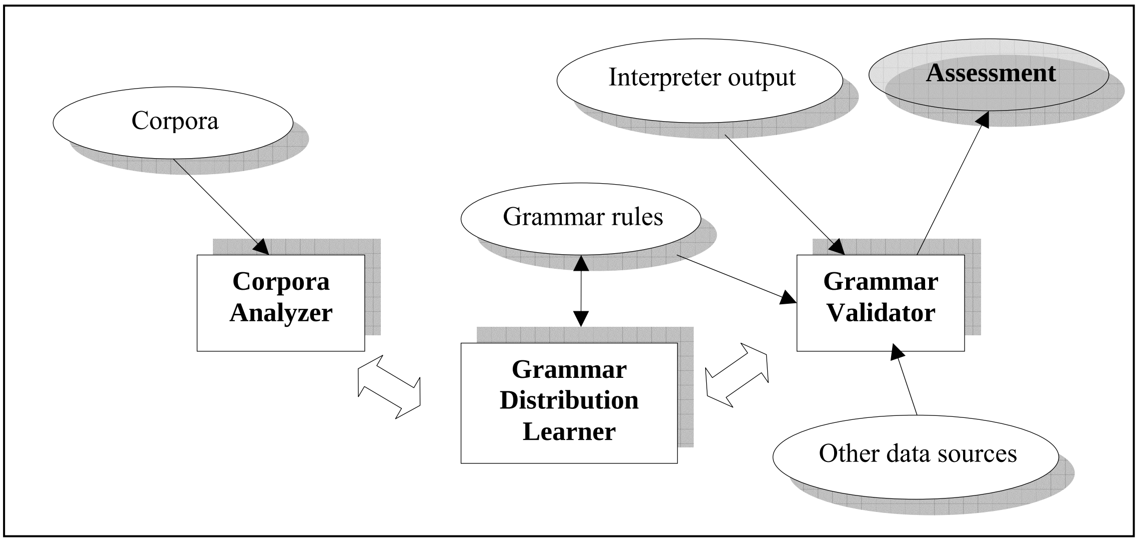


**Fig. 1** Components design for the grammatical assessment tool

### 2.1.Corpora Analyzer

This component is in charge of performing the statistical analysis of the linguistically tagged corpora. According to a basic set of morphological and syntactic labels (taken from TreeBank [1]), the most frequent combinations of morphosyntactic structures are extracted. This frequency value is the base of the set of the statistics to be computed. The output of this module is the set of grammar rules that are going to be contrasted with the text generated by the interpreter. Each rule will also be accompanied by a relevance factor, marking the influence of failure of the interpreter in following this rule. The corpora used for training consists of different sets of tagged English texts taken from newspapers but, in a future approach, the tool can be used with any other corpora available and with the corresponding linguistic label set. The linguistic analysis for the input text is done using the Cherniak's parser[1] for English [3]. In Fig. 2 is included an example of the XML tagging mechanism used to store the analysis of the sentence "*All of the GM divisions except Cadillac showed big declines*".


```
<?xml version="1.0"?>
<sentence id="0">
<morphosynt>
  <syntelem type="S1" id="0S1">
      <syntelem type="S" id="0S2">
        <syntelem type="NP" id="0S3">
          <syntelem type="NP" id="0S4">
            <word id="0W1" pos="DT">All</word>
          </syntelem>
          <syntelem type="PP" id="0S5">
            <word id="0W2" pos="IN">of</word>
```


[1] http://www.cs.brown.edu/people/ec/#software

```
          <syntelem type="NP" id="0S6">
            <word id="0W3" pos="DT">the</word>
            <word id="0W4" pos="NNP">GM</word>
            <word id="0W5" pos="NNS">divisions</word>
          </syntelem>
        </syntelem>
        <syntelem type="PP" id="0S7">
          <word id="0W6" pos="IN">except</word>
          <syntelem type="NP" id="0S8">
            <word id="0W7" pos="NNP">Cadillac</word>
          </syntelem>
        </syntelem>
      </syntelem>
      <syntelem type="VP" id="0S9">
        <word id="0W8" pos="VBD">showed</word>
        <syntelem type="NP" id="0S10">
          <word id="0W9" pos="JJ">big</word>
          <word id="0W10" pos="NNS">declines</word>
        </syntelem>
      </syntelem>
    </syntelem>
  </syntelem>
</morphosynt>
</sentence>
```


**Fig. 2** Result of the linguistic analysis for the sentence "*All of the GM divisions except Cadillac showed big declines*" expressed in XML

### 2.2.Grammar Distribution Learner

This component is included to allow a dynamic development of the set of grammatical rules to be contrasted with the interpreter-generated text. Taking input from the Grammar Validator component (which collects user interactions), the Grammar Distribution Learner stores information about the validity of the rules and about the type of mistakes or inappropriate translations that interpreters make more often.
There are two main kinds of rules used to represent the content of the analysed text. Both rule types take the verb as the central element and are:

- *Morphological rules*: The part-of-speech tags around the verb are taken. A predefined radius, measured in number of elements, is used to limit the number of tags to be considered taking the verb as the initial element.
- *Word based rules*: The words around the verb are taken. Like in the morphological rules, a predefined radius, measured in number of elements, is used to limit the number of neighbour words to be taken.

In Fig. 3 a screenshot of the grammar distribution learner results is shown.

### 2.3.Grammar Validator

This is the core module of the tool, where the grammatical assessment process is carried out. Taking as input the text generated by the interpreter and a set of grammar rules to be contrasted, the Grammar Validator generates the quality assessment for that interpreter. A weighting formula is used to compute the degree of adaptation of the text produced by the interpreter. This formula is shown bellow:

$$\begin{aligned} \mathit{Weight} = & \sum(\mathrm{Freqexact}) * \mathrm{Kexact} + \sum(\mathrm{Freqpartial}/\mathrm{Npartial}) * \mathrm{Kpartial} + \\ & \sum(\mathrm{FreqpartialLeft}/\mathrm{NpartialLeft}) * \mathrm{KpartialLeft} + \\ & \sum(\mathrm{FreqpartialRight}/\mathrm{NparrtialRight}) * \mathrm{KpartialRight} \end{aligned} \quad (1)$$


**Fig. 3** Results of the Grammar Validator component

where *Kexact*, *Kpartial*, *KpartialLeft* and *KpartialRight* are constants that can be changed according to the quality of the assessments. They are defined as configuration parameters of the system.
The set of rules can be semi-automatically built using the Corpora Analyzer component, but it can also be supplied as an external input. The only requisite to be fulfilled is the format in which these rules must be provided. This format is defined using the XML language.

All these components have been developed in a web environment, using JAVA® technology [4], a Tomcat[2] servlet engine and implemented as a Web Service to allow an easier integration with the rest of the components of the ASSESSTI platform.

## 3. Conclusions

There are several advantages related to the selected approach to the non-typical sentences detection module presented here that should be pointed out. First of all, it is easily adaptable to languages for which parsers for linguistic analysis are available. This includes almost all European languages. Having this analysis tool available, the rest of the applied methods are statistical in nature; so no language-specific knowledge is needed. Secondly, the production of rules can be based on several corpora, allowing a wider range of sources to train the system. Thirdly, it is possible to include external resources available for a particular language. Due to the verb-centered approach taken, information contained in theaurus like FrameNet [2] or VerbNet [5] can be applied to the assessment process. Finally, the technological environment selected, using the JAVA® platform and a Web Service model, assures an easy

[2] http://tomcat.apache.org

integration with the rest of the components of the ASSESSTI platform and also a portable component for other systems where linguistic structures should be compared.

## References


[1] B. Santorini, "*Part-of-speech tagging guidelines for the Penn Treebank Project (3rd revision*)," Department of Computer and Information Science, University of Pennsylvania, Philadelphia, Tech. Rep. MS-CIS90 -47, Line Lab 178, (1990), ftp://ftp.cis.upenn.edu/pub/treebank/doc/ manual/root.ps.gz. 89

[2] C. J. Fillmore, C. F. Baker and H. Sato, H. "*The FrameNet Database and Software Tools"*. In Proceedings of the Third International Conference on Language Resources and Evaluation (LREC), (2002) pp.1157-1160.

[3] E. Charniak, "*Statistical parsing with a context-free grammar and word statistics"*, Proceedings of the Fourteenth National Conference on Artificial Intelligence AAAI Press/MIT Press, Menlo Park (1997).

[4] G. Bracha, J. Gosling, B. Joy, and G. Steel. "*The Java Language Specification*". Addison Wesley, second edition, June 2000.

[5] K. Kipper, H. T. Dang and M. Palmer, "*Class-Based Construction of a Verb Lexicon"*. AAAI-2000, (2000)