Text difficulty evaluation approaches should consider both vocabulary and grammar. Features used in text difficulty evaluation for L2 learners should derive from what students actually learn from language lessons. Text difficulty evaluation studies and educational technology researches for language learning should pay more attention to language-specific linguistics.

Grammatical Templates

We introduce grammatical templates as features for text difficulty evaluation because:

- They are units of grammar that expert language instructors and linguists have identified as the most important grammatical knowledge for a specific language.
- They are typically emphasized as key points in every textbook lesson.
- Students (usually explicitly) learn grammatical templates in language lessons.
- We believe they reflect the conceptual units of grammar more closely than parse trees.

**Translation**

<table>
<thead>
<tr>
<th>Template</th>
<th>Pronunciation</th>
<th>Translation</th>
<th>Notation</th>
</tr>
</thead>
<tbody>
<tr>
<td>~ は ~ の ~ を ~ ない</td>
<td>~ wa ~ no ~ wo ~ wa ~ ni</td>
<td>&quot;he will soon arrive in Tokyo&quot;</td>
<td>A B C D E F</td>
</tr>
<tr>
<td>~ と ~ の ~ を ~ ない</td>
<td>~ to ~ no ~ wo ~ wa ~ ni</td>
<td>&quot;I go to see my girlfriend&quot;</td>
<td>A B C D E F</td>
</tr>
<tr>
<td>~ は ~ と ~ の ~ を ~ ない</td>
<td>~ wa ~ to ~ no ~ wo ~ wa ~ ni</td>
<td>&quot;this is not your book&quot;</td>
<td>A B C D E F</td>
</tr>
</tbody>
</table>

Grammatical Templates in Japanese

Grammatical templates play an important role in language understanding because:

- Many grammatical templates suggest sentence structure.
  - English: "hardly ... when ..."
  - German: "nicht nur ..., sondern auch ..." (not only ... but also ...)
  - Japanese: "必ずしも ... とはいええない" (it is not necessarily the case that ...)
- For languages like Chinese and Japanese, lacking knowledge of some grammatical templates will cause difficulties in segmentation.
  - Japanese: "...と ..." (two opposite behaviors occurring alternately) in the phrase "行きつづくり" (to walk back and forth)
  - Chinese: "越-越好" (the more the better) in "越早越好" (the earlier the better)
- Some grammatical templates may refer to special meanings that cannot be understood as the combination of individual words.
  - English: "in terms of", "such that"
  - German: "mit etwas zu tun haben" (have something to do with ...)
  - Japanese: "... こたない" (no need to ...)

**Approach**

- First, we build a library of grammatical templates. Templates are labeled with difficulty.
  - We use JLPT [1] levels: from N5(easy) to N1(hard).
- Given a sentence, our algorithm then searches through each node of its parse tree and identifies grammatical templates by matching the phrase associated with the children of that node to all templates in the library.
- For each text, we calculate the per-sentence numbers of grammatical templates at each difficulty level, and use them as features to predict text difficulty.

To predict the difficulty levels of a given text, we used not only the conventional machine learning algorithms like kNN and SVM, but also a simple and novel algorithm.

**Multilevel Linear Classification (MLC)**

We notice that texts of adjacent levels are nearly linear-separable with two features: templates "on the boundary" and templates "outside the boundary".

- Taking advantage of this phenomenon, we build 4 linear classifiers for 4 pairs of adjacent levels. Our algorithm works as follows:
  - A text is judged by the N5 classifier first. If it is no harder than N5, it will be labeled as an N5 text; otherwise, it will be passed to the N4 classifier in order to decide if it is harder than N4.
  - The process continues similarly, until it is judged to be harder than N2, it will be labeled as an N1 text.

**Results**

- **Grammatical Template Features:** Average number of N1-N5 grammatical templates per sentence
- **Baseline Readability Features:** Features used in recent readability evaluation studies
- **Hybrid Features:**
  - Grammatical Template Features + Baseline Readability Features

<table>
<thead>
<tr>
<th>Feature Set (number of features)</th>
<th>Algorithm</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>TF-IDF Features (5100)</td>
<td>kNN</td>
<td>69.10%</td>
</tr>
<tr>
<td></td>
<td>SVM</td>
<td>80.50%</td>
</tr>
<tr>
<td>Baseline Readability Features (12)</td>
<td>kNN</td>
<td>72.30%</td>
</tr>
<tr>
<td></td>
<td>SVM</td>
<td>80.90%</td>
</tr>
<tr>
<td>Grammatical Template Features (5)</td>
<td>kNN</td>
<td>78.00%</td>
</tr>
<tr>
<td></td>
<td>SVM</td>
<td>81.10%</td>
</tr>
<tr>
<td></td>
<td>MLC</td>
<td>87.70%</td>
</tr>
<tr>
<td>Hybrid Features (17)</td>
<td>kNN</td>
<td>85.70%</td>
</tr>
<tr>
<td></td>
<td>SVM</td>
<td>88.50%</td>
</tr>
</tbody>
</table>

**Conclusion**

- Adding grammatical template features to existing readability techniques significantly improves text difficulty evaluation for language learners.
- The Multilevel Linear Classification algorithm, which uses grammatical template features only, has high accuracy, a small number of features, and a simple, human-understandable structure.

**References**

[1]. Japanese-Language Proficiency Test: www.jlpt.jp