

Challenges Faced During Annotation

The annotation of a learner corpus is a complex endeavor that involves not only the technical task of labeling errors but also the conceptual and organizational challenges of defining categories, ensuring consistency, and coordinating multiple annotators. Inevitably, a number of challenges emerged throughout the process, and the strategies developed to address them became an integral part of shaping a coherent and reliable annotation.

Before the annotation process, a set of error types had already been established by an error-tagged learner corpus (Golynskaia, 2022). These error types were subsequently revised in the preparatory stage, when two expert annotators undertook the task of updating the list in line with the principles of the faceted taxonomy for classifying errors in language use.

The first challenge faced at this stage concerned the lack of precise descriptions for certain categories in the existing literature on Teaching Turkish as a Foreign Language (TTFL), both in terms of form and content as well as in naming conventions. A case in point was the vowel appearing before the present tense affix *-yor* when it attaches to a verb root ending in a consonant. While some scholars classify these vowels as part of the allomorph *-yor/-Iyor*, others regard them as epenthetic. However, the literature on vowel epenthesis in Turkish rarely, if ever, includes examples involving the present tense. This ambiguity naturally led to differences among annotators. It was therefore necessary to reach a common decision and establish a clear stance for naming such errors. In the end, such cases were systematically evaluated within the TTFL context, and appropriate tags were developed to capture them. These tags, along with their defined scope, were incorporated into the guideline, providing annotators with a consistent reference and reducing the risk of divergent interpretations.

Once the preparatory challenges had been resolved, the annotation phase brought a new set of difficulties. A recurrent problem, which was related to the Label Studio, was the persistent pre-selection of labels. When annotators failed to deactivate labels after use by pressing the “Esc” key, the previously selected label remained active, leading to unintended label application on subsequent annotations, even when a different label was intended. To minimize such issues, annotators received training prior to the annotation phase. However, minor technical problems of this kind inevitably persisted. This particular issue was later resolved by reviewing the annotations, which helped reduce annotation inconsistencies to a minimum.

A further challenge concerned punctuation, arising from the limitations and incompleteness of the punctuation guidelines provided by the Turkish Language Association (*Türk Dil Kurumu, TDK*), the official regulatory body for the Turkish language. While the TDK website offers specific rules (for example, explicitly prohibiting commas before or after the conjunctions “and”, “or”, “either...or...”, or repeated conjunctions), these guidelines prove insufficient in their coverage. Notably, they lack explicit direction regarding the use of commas with other conjunctions, such as “but”. Consequently, in the absence of comprehensive and standardized rules, annotators were compelled to rely on their individual interpretations and pre-existing habits when annotating

punctuation errors, potentially introducing inconsistencies into the annotation process. Its impact was mitigated by adopting the TDK rules as the primary reference and by acknowledging that certain unresolved gaps would inevitably result in minor variations among annotators, regarded as a natural part of the process.

Another challenge arose from the difficulty in reliably determining the nature of certain errors, as they could be categorized in multiple ways. One particular case that required careful consideration was distinguishing between diacritic errors and vowel harmony errors (e.g., “*evim*” (my house) (from this point onward, the translations correspond to the correct forms of the erroneous samples), “*yüzüğü*” (the ring)). In some instances, this distinction was not clear enough. As a result, annotators decided that in ambiguous cases, both labels should be applied. This dual-labeling approach provided more comprehensive feedback to the learner and ensured that all possible explanations were represented. Another example involved the challenge of distinguishing between redundant use of the “-i/-i/-u/-ii” affix as either a possessive or accusative marker (e.g., *Bu kitabı ilginç.*” (This book is interesting.)). To resolve this, annotators applied the “Unnecessary Affix” label, thus prioritizing methodological consistency over context-specific granularity.

Yet another challenge involved distinguishing between primary and secondary errors. In many cases, an initial error, such as a spelling, triggered subsequent grammatical errors, resulting in multiple visible mistakes in the same segment. Annotators needed to identify and tag the primary error, recognizing that the secondary errors were a consequence of it. This can be clarified by examining the word “*yuzuklar*” (rings) as an example. The primary error here is the omission of diacritical marks, which results in an incorrect spelling of the word. Once this primary error is corrected restoring the proper spelling “*yüzük*”, it becomes clear that the use of the plural affix *-lar* is inappropriate in this case, since vowel harmony dictates the use of its *-ler* counterpart. However, this secondary error should not be tagged, because it directly stems from the primary one and would not have occurred independently. This required a high level of attention and careful judgment during annotation. Distinguishing between primary and secondary errors was essential for maintaining the accuracy and reliability of the annotated data. Without this distinction, error counts could be inflated by tagging multiple visible errors that were in fact consequences of a single underlying error. To support consistent decision-making, additional explanatory notes were incorporated into the guideline, providing annotators with clearer instructions on how to handle such cases.