PLENE WRITING

Another example of a feature to distinguish in a letter is the usage of plene vs. non-plene writings. A plene writing is an extra sign written to highlight the vowel of a previous CV-sign or the next VC-sign. The two signs would thus be CV-V or V-VC where the vowels are equivalent. Such a writing is considered “extra” or plene because it is not strictly necessary in a syllabic writing system, but is nevertheless used, usually when a vowel is long. The distribution of the usage of plene writings vis a vis their oppositional counterparts (CV-V or CV) could, like other features, indicate a particular scribe’s preference.

One reason to consider plene writings as a feature for clustering is a comment made by Parpola for LAS 14 (1983:19). There, he proposes the possibility that this letter was written by Akkullānu given his preference in other letters to write in a plene style. If clustering based on this feature yields significant results, it may validate this argument for authorship of a letter.

However, one must be careful when applying this argument across any oppositional pair. Certain words have a preference across the corpus to be written one way instead of the other. For example, the word *mā*, a particle introducing direct speech, has a major tendency (96% of the time) to be written in the plene form <ma-a> as opposed to <ma>. In contrast the words *lā*, a negative particle, and *lū*, a precative particle, have a bend toward the non-plene writing (84% and 73%). Therefore, the plene writings of <ma-a>, <la-a>, and <lu-u> should not be combined into one feature because if even if a scribe prefers to write without extra V-signs, they may still write <ma-a> due its heavy expectation. It may be possible to get around this problem by excluding words like *mā*. (This remains to be done)

PLENE WRITING (coding)

The method of finding plene writings in corpus requires some work. I composed a regular expression to select the word forms that contain a V-sign copy of the previous or following sign’s vowel and filtered the data frame on the *form* column. This only gave me the possible words (lemma, normalization, and form) that could have a plene writing. The next step was to find their non-plene counterparts. (more needed here)