

DaRomajiBearbLesung

Enthält die Lesung in Hiragana mit Markierungen zur Segmentierung der Einträge durch Markierungen, Leerzeichen, Unterstriche, Tilden, Nakaten, Pipes etc.. Mit diesem Mark-up kann man verschiedene Geschmacksrichtungen Latein-Umschrift und Furigana generieren.

Bei Untereinträgen stehen die Kana des entsprechenden Haupteintrags in Spitzklammern und könnten für verkürzte Schreibweise durch Tilde ersetzt werden. Wenn der Haupteintrag aufgrund von z.B. Konjugation nicht wirklich mit dem Haupteintrag identisch ist, ist der öffnenden Spitzklammer ein Schrägstrich "/" als Escape-Zeichen vorangestellt. Wenn sich ein Untereintrag auf mehrere Haupteinträge bezieht, sind die Spitzklammern nummeriert.

Das Feld dient dazu, die Hiragana z.B. per Formel in Romaji umzurechnen.

Markiert werden Joshi, bei denen Umschrift problematisch ist (ist wa nicht ha und ist e nicht he), mit "Jo".

Bei Großschreibung wird "Gr" voran gestellt.

Devokalisierung wird mit "Dev" vor der entsprechenden Silbe angezeigt. Bei diesen Silben könnte bei Lateinumschrift ein kleiner Kringel unter i, u oder o gestellt werden.

G-Laut im Wortinneren, der nicht nasalisiert wird, ist mit "NN" vor der Silbe markiert; in der Lateinumschrift könnte man z.B. einen Apostroph vor diese Silbe setzen. Das betrifft Fremdwörter, bei denen der velare Verschluss nicht nasalisiert ist.

Leerzeichen stehen vor Worten.

Tilden "~" stehen vor dem flektierten Teil eines Verbs oder Verbaladjektives.

Nakaten steht, um Worte weiter zu segmentieren und lesbarer zu machen.

Pipe steht bei Segmentierung, die für Furigana eine Rolle spielt aber nicht für Lateinumschrift.

Unterstriche stehen, wenn eine Silbenkombination zu einer Längung führen könnte, obwohl zwei Vokale stehen sollen. Möglicherweise überschneidet sich das mit Verwendung der Apostrophen und fällt irgendwann weg.

Apostroph trennt Silben, die sich auf unterschiedliche Kanji beziehen; das wäre wichtig für Furigana.

Um DIN-Umschrift zu erreichen, müssten einige dieser Nakaten in Leerzeichen umgewandelt werden und einige müssten ignoriert werden. An manchen Stellen würde ich Leerzeichen setzen, wo DIN keines möchte (z.B. Nomen mit suru). Hierfür müssen wir uns noch weiteren Markup ausdenken.