**Controle Nederlandse lemma’s van Groningse woorden**

De kolom token\_GN bevat Groningse tekst. In de kolom staan de woorden onder elkaar. De kolom lemma\_NL bevat Nederlandse lemma’s die geautomatiseerd zijn gegenereerd. Daardoor zijn de lemma’s niet altijd correct. De bedoeling is om de foute lemma’s te corrigeren. Hieronder volgen wat voorbeelden van gevallen waar de keuze van het Nederlandse lemma misschien niet vanzelfsprekend is.

* Vertaal elk Gronings woord met het Nederlandse woord dat er het meest op lijkt. Voorbeelden:

|  |  |  |
| --- | --- | --- |
| **token\_GN** | **lemma\_NL** |  |
| hoes | huis | i.p.v. woning |
| gewoon | gewoon | i.p.v. gewend |
| luu | lui | i.p.v. mensen |

* Let op sommige tijdsaanduidingen. Voorbeelden:

|  |  |  |
| --- | --- | --- |
| **token\_GN** | **lemma\_NL** | **betekenis** |
| smörgens | morgens | ‘s morgens |
| smiddags | middags | ‘s middags |
| soavends | avonds | ‘s avonds |
| voorjaars | voorjaars | in het voorjaar |
| haarstmis | herfsts | in de herst |

* Bij plaatsnamen wordt soms een lidwoord toegevoegd, terwijl dat in het Nederlands niet zo is. Voorbeelden:

|  |  |  |
| --- | --- | --- |
| **token\_GN** | **lemma\_NL** | **upos** |
| n | een | DET |
| Daam | Appingedam | PROPN |
|  |  |  |
| de | de | DET |
| Penne | Opende | PROPN |

* Vertaal hulpwerkwoorden letterlijk (maar wel als infinitief), ook als daardoor de Nederlandse vertaling wat vreemd aandoet. Voorbeelden:

|  |  |  |
| --- | --- | --- |
| **token\_GN** | **lemma\_NL** | **betekenis** |
| (Dat) het (goud west) | hebben | i.p.v. zijn |

* Handhaaf altijd de lidwoorden, ook als dit in het Nederlands niet correct is. Voorbeeld:

|  |  |  |
| --- | --- | --- |
| **token\_GN** | **lemma\_NL** | **betekenis** |
| de (streektoalkrais) | de (streektaalgebied) |  |

Let op gevallen van suppletie, zie: <https://nl.wikipedia.org/wiki/Suppletie_(taalkunde>).

Voorbeelden:

Het lemma van ‘bent’, ‘wezen’ is ‘zijn.

Het lemma van ‘beter’, ‘best’ is ‘goed’.