

| language | dataset | parent dataset | Skipgram | fastText, h=10,<br>bigram | MDLText |
|----------|---------|----------------|----------|---------------------------|---------|
|          |         | average        | 57,40%   | 0,00%                     | 0,00%   |
| AR       | WS353   | -              | 51,00%   |                           |         |
| DE       | GUR350  | -              | 61,00%   |                           |         |
| DE       | GUR65   | -              | 78,00%   |                           |         |
| DE       | ZG222   | -              | 35,00%   |                           |         |
| EN       | RW      | -              | 43,00%   |                           |         |
| EN       | WS353   | -              | 72,00%   |                           |         |
| ES       | WS353   | -              | 57,00%   |                           |         |
| FR       | RG65    | -              | 70,00%   |                           |         |
| RO       | WS353   | -              | 48,00%   |                           |         |
| RU       | HJ      | -              | 59,00%   |                           |         |