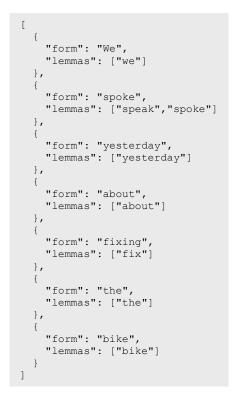# bitext

Bitext DLA Platform: Lemmatization

Lemmatization provides all potential roots (also known as *lemmas*) for a given word, using morphological analysis and carefully curated lexicons. For example, for the word **"spoke"**, the lemmatization service will return:

```
[
  {
    "form": "spoke",
    "lemmas": ["speak", "spoke"]
  }
]
```

For the sentence **"*We spoke yesterday about fixing the bike*"**, the lemmatization service will output the following data:

```
[
  {
    "form": "We",
    "lemmas": ["we"]
  },
  {
    "form": "spoke",
    "lemmas": ["speak","spoke"]
  },
  {
    "form": "yesterday",
    "lemmas": ["yesterday"]
  },
  {
    "form": "about",
    "lemmas": ["about"]
  },
  {
    "form": "fixing",
    "lemmas": ["fix"]
  },
  {
    "form": "the",
    "lemmas": ["the"]
  },
  {
    "form": "bike",
    "lemmas": ["bike"]
  }
]
```

Spoke: past tense of the verb to speak
Spoke: noun