TextCritical.net - Bug #2611

UnicodeDecodeError: 'ascii' codec can't decode byte 0x80 in position 4: ordinal not in range(128)

03/06/2020 06:09 PM - Luke Murphey

Status:	Closed		**Start date:**	03/06/2020
Priority:	Normal		**Due date:**	
Assignee:			**% Done:**	100%
Category:			**Estimated time:**	0.00 hour
Target version:	3.2			

Description
Not Found: /api/search/None

History

#1 - 03/06/2020 06:20 PM - Luke Murphey

```
Performing a search, page=1, page_len=10, include_related_forms=False, search_query="None "
Search query parsed, default_search_field="content", raw_query="content:<none>"
Looking for variations of the search term in order to perform a search, word=none
```

#2 - 03/09/2020 07:14 PM - Luke Murphey

Qs:

- **No:** Does the issue repro locally?
  - No, at least not with a fresh index
  - http://localhost:8080/search?q=work%3Aabdicatus+section%3A%221%22+%CF%80%CE%B1%CF%81%E1%BD%B1%CE%B4%CE%BF%CE%BE%CE%B1+
  - http://localhost:8080/search?q=work%3Acontra-apionem+section%3A%221.1%22+%CF%80%CF%81%CE%BF%CF%83%E1%BD%B3%CF%87%CE%BF%CE%BD%CF%84%CE%B1%CF%82
- **Yes:** What happens if I reindex two works, will the indexes be persisted?
  - Yes
- **No:** Does re-indexing a work fix this for that one work?
  - backed up indexes using "sudo tar -czvf indexes.tar.gz indexes/"
  - Got a 404 when reindexing
  - The indexes did get changed in that some new files got created
  - Error is no different
- Are others having this issue?
  - https://github.com/WeblateOrg/weblate/issues/1508

Obs:

- Get an error when trying to index a work on the prod server

Solns:

- Handle the content as Unicode in UTF-8
- The indexes need to be recreated
  - Recreating one with "python manage.py make_search_indexes -c -w against-apion"
  - That worked
  - docker exec -it pedantic_goldberg python manage.py make_search_indexes -c

#3 - 03/09/2020 10:55 PM - Luke Murphey

- Status changed from New to Closed

- % Done changed from 0 to 100