

cub[AI]^x

NeuroSpell

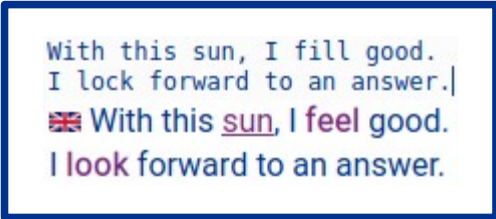
The only fully automatic corrector for spelling and grammar,
based on neural technology (Deep Learning)

<http://neurospell.com>



NeuroSpell

- Designed for automatic correction decisions
 - ➔ tunable decision threshold
- Based on a neural statistical model
 - ➔ trainable on specific in-domain data
 - ➔ trainable on any language
- Speech-To-Text or OCR errors may be very different from human errors
 - ➔ adaptable for specific error types
 - ➔ can produce corrections on grammatically correct sentences
- Rewrites all words of a sentence taking each sentence as a whole
 - ➔ can correct errors with far dependencies in a sentence
 - ➔ unchanged words are validated



With this sun, I fill good.
I lock forward to an answer.
🇬🇧 With this sun, I **feel** good.
I **look** forward to an answer.

Checking is not correcting



With **this sun**, I fill **good**.
I lock forward to an answer.

thissun
These words should be merged.



With this sun, I **fill** good. I lock forward to an answer.

• SPELLING

fill → **feel**

The word **fill** doesn't seem to fit this context.
Consider replacing it with a different one.



LanguageTool
Votre assistant d'écriture

With this sun, I fill good.

I lock forward to an answer.

✓ Tout a l'air bon. Aucune erreur trouvée.



With this song, I fill good. I lock forward to an answer.



ReversoCorrecteur

With this song, I fill good.
I lock forward to an answer.



With this sun, I fill good. I lock forward to an answer.

fill → feel lock → am looking



SpellCheckPlus.com

With this sun, I **fill good**. I lock forward to an answer.

You should probably use the adverbial form **well**, e.g.: *She works **well**.*

Rule-based systems are designed to find doubtful words.
They may provide with suggestions to be decided on manually.

Rule-based checkers

- 1) Grammar analysis (verb, noun, adjective, etc.)
- 2) Lexical and grammar inconsistencies detection (dictionaries, hand-written rules, statistic models)

- PRO:

- Often explainable

- CONS:

- Hard to use for automatic correction (checking is not correcting / doubt vs choice)
- May be confused by errors or unknown words
The grammar analysis may be wrong even on good sentences
- Hardly handle far dependencies in a sentence
- Hardly find errors providing with a grammatically correct sentence
- Untagged/unchanged words are not validated
- Best ones are not available in a lot of languages
- Hard to adapt to domain-specific vocabulary and phrases
- Not intended to handle machine-specific speech to text errors (built for human errors)

Take a **free** tour
to **free** your mind.

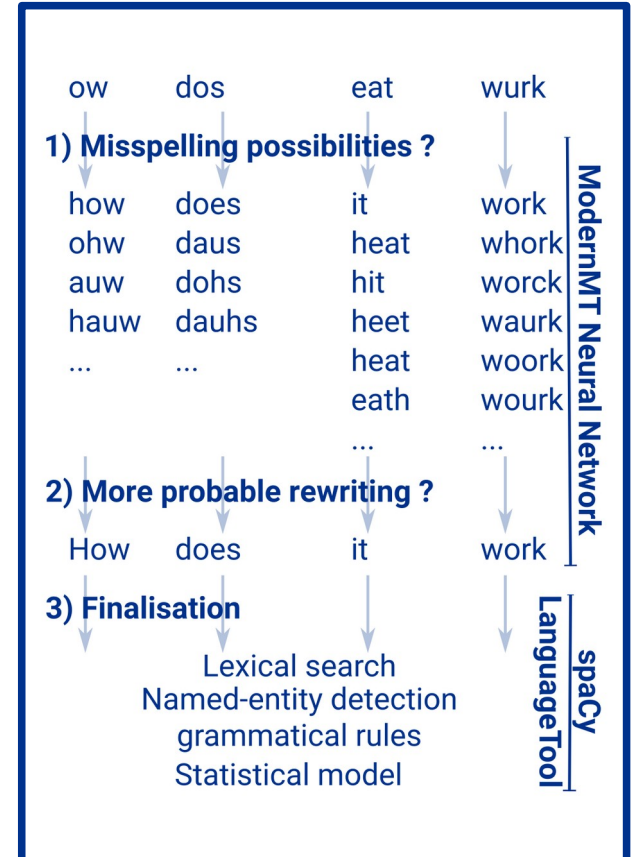
Take (VERB)
a (DET)
free (ADJ)
tour (NOUN)
to (PART)
free (VB)
your (PRON)
mind (NOUN)
. (PUNCT)



NeuroSpell

- 1) Evaluates all ways to re-write each word or set of words of a sentence (possible variations need to be defined or collected)
- 2) Produces the most statistically probable written form taking into account the statistical plausibility of the sentence to be corrected (tunable balance threshold)

- PRO:
 - Designed to really perform automatic corrections
 - Produced quality is (nearly) insensitive to pre-defined errors (writing variations) whatever their nature and their number in a sentence
 - Can handle far dependencies in a sentence (sentence analysed as a whole)
 - Can correct errors producing grammatically correct sentences (statistical effect)
 - Unchanged words are validated (according to possible writing variants)
 - Can be built for any new language at a reasonable cost (currently FR + EN)
 - Trainable on domain-specific vocabulary or phrases (large training data needed)
 - Can be trained to correct very machine-specific speech to text errors
 - Is combined with a rule-base checker (hybrid system)
- CONS:
 - Often not explainable



cub[AI]^x

Etienne Monneret

- 10 years: multi-lingual document industry (Jouve)
- 14 years: translation software edition (Lingua et Machina)
- 2 years: Linguistic AI expert (freelance) / software developer
 - Unique neural-based automatic spell and grammar corrector
 - Automatic terminology extraction
 - Semantic pairing / indexing
 - Domain-adapted Neural Machine Translation (NMT)
 - Translation workflow / CAT tool / Translation Memory
 - ... multi-lingual document comparison, stereoscopic video editor, etc.

<http://cubaix.com>

